Does the Dark Triad manifest similarly in men and women?

Chiorri, Carlo; Garofalo, C.; Velotti, Patrizia

Published in:
Current Psychology

DOI:
10.1007/s12144-017-9641-5

Publication date:
2019

Document Version
Early version, also known as pre-print

Chiorri, C., Garofalo, C. & Velotti, P. (2017). Does the Dark Triad Manifest Similarly in Men and Women? Measurement Invariance of the Dirty Dozen across Sex. *Current Psychology*, Advanced Online Publication, doi: 10.1007/s12144-017-9641-5

RUNNING HEAD: ITALIAN VERSION OF THE DD

Does the Dark Triad Manifest Similarly in Men and Women? Measurement Invariance of the Dirty Dozen across Sex

Carlo Chiorri[1], Carlo Garofalo[2], Patrizia Velotti[1]

[1] Department of Educational Sciences, University of Genova, Corso A. Podestà, 2, 16128 Genova (Italy)

[2] Department of Developmental Psychology, Tilburg University, PO Box 90153, 5000 LE Tilburg (The Netherlands)

Conflict of Interest: On behalf of all authors, the corresponding author states that there is no conflict of interest.

Correspondence to:
Carlo Chiorri, PhD
Department of Educational Sciences
University of Genova
Corso A. Podestà, 2
16128 Genova (Italy)
E-mail: carlo.chiorri@unige.it
Tel:+39 010 20953709
Fax: +39 010 20953728

Abstract

The Dark Triad is a constellation of three socially undesirable personality traits: narcissism, psychopathy, and Machiavellianism. Previous research has shown that men tend to score higher than women on Dark Triad scales, but the validity of these results is questionable as there is no evidence that the scales used exhibit measurement invariance across sex in the adult population. Here, we report four studies assessing the measurement invariance across sex of a recently developed, concise measure of the Dark Triad, namely Jonason and Webster's (2010) Dirty Dozen (DD). As no validated Italian version of the DD was available, we developed an Italian version and assessed its psychometric properties. Studies 1 to 3 revealed that the Italian DD had adequate psychometric properties, and replicated the three-factor structure and the nomological network of the original version. Study 4 provided evidence of the measurement invariance of the DD across sex, such that men scored higher than women with respect to psychopathy, Machiavellianism, and, to a lesser extent, narcissism. These findings indicate that the DD can be used to provide reliable assessments of sex differences in Dark Triad traits. Furthermore, the results of sex comparisons are consistent with a biosocial approach to social role theory that assumes that being agentic rather than communal is considered desirable for men and undesirable for women.

Keywords: Dark Triad; measurement invariance; sex differences; Machiavellianism; psychopathy; narcissism

Does the Dark Triad Manifest Similarly in Men and Women? Measurement Invariance of the Dirty Dozen across Sex

Psychopathy, narcissism, and Machiavellianism are three distinct, albeit overlapping, personality constructs that are said to represent the dark side of human nature and are often studied together under the label 'Dark Triad' (DT; Paulhus & Williams, 2002). Psychopathy can be defined as an antagonistic disposition characterized by deceptiveness, pathological lying, the absence of remorse and guilt, and tendencies to irresponsible and impulsive behavior (Hare & Neumann, 2008). Narcissism is defined by a pattern of grandiosity, an inflated sense of self, a sense of entitlement, dominance, exhibitionism, and superiority (Morf & Rhodewalt, 2001). A tendency to interpersonal exploitation and callousness is common to narcissism and psychopathy. However, there is substantial evidence that narcissism also encompasses characteristics indicative of vulnerability, such as fragile or contingent self-esteem, emotion dysregulation, hypersensitivity to rejection, and consequent social avoidance (Cain, Pincus, & Ansell, 2008). The vulnerable side of narcissism also includes negative feelings such as helplessness, emptiness, and shame (Velotti, Elison, & Garofalo, 2014). Finally, Machiavellianism is defined as a duplicitous interpersonal style accompanied by cynicism and a pragmatic approach to morality (Jones & Paulhus, 2009). Machiavellian individuals prioritize money, power, and competition over community building, self-love, and family concerns, and they use manipulative interpersonal strategies, such as flattery and lying, to achieve their goals (Jones & Paulhus, 2009).

Although the DT traits share a common core of disagreeableness, seminal studies have shown that, at least in non-clinical populations, they are sufficiently distinct to warrant separate assessment (Paulhus & Williams, 2002). This is reflected in the typically moderate correlations among standard measures of the DT (Furnham, Richards, & Paulhus, 2013). In line with a dimensional approach to personality assessment (Wiggins & Pincus, 1989), the pathological forms of these 'dark' personality traits are considered extremes of normal personality traits. It is therefore

important to extend their investigation from *clinical* samples to *subclinical* samples (i.e., broader community samples; Furnham et al., 2013).

Sex Differences in the Dark Triad

One of the most consistent findings on DT personality traits relates to sex differences in narcissism (Grijalva et al., 2015) and psychopathy (Cale & Lilienfeld, 2002), such that men score higher than women on measures of both traits, while the picture of sex differences in Machiavellianism is much less consistent (Furnham et al., 2013). Wood and Eagly's (2012) biosocial construction model has been used to account for sex differences in the DT (e.g., Grijalva et al., 2015). This model posits that the biological specialization of the sexes resulted in a gendered division of labor in traditional societies, and that this division of labor gave rise to sex role beliefs and stereotypes. Most sex stereotypes can be categorized as agentic (e.g., competitiveness, dominance, and need for achievement) or communal (e.g. friendliness, nurturance, and selflessness) characteristics. Communal characteristics are compatible with the social roles derived from women's reproductive activities, whereas agentic characteristics are consistent with the social roles linked to men's traditional access to wealth-generating activities. Many of the correlates of DT traits seem to reflect agentic rather than communal characteristics (Jones & Paulhus, 2010), and being agentic rather than communal is considered desirable for men and undesirable for women (Rudman, Moss-Racusin, Phelan, & Nauts, 2012).

One potentially problematic feature of comparisons of DT scores across sex is that evidence of the measurement invariance across sex of the scales used is not always available. Unless the DT scales are measuring the same construct in the same way in both sexes, and the scales operate in the same way in both sexes, comparisons of manifest means are likely to be invalid (Millsap, 2011). Moreover, an apparent difference in latent means might disappear if results are corrected for lack of measurement invariance. To the best of our knowledge, only one study has addressed this issue (Klimstra, Sijtsema, Henrichs, & Cima, 2014), using the Dirty Dozen (DD, Jonason & Webster, 2010) as a measure of the DT. The DD is a concise questionnaire that assesses the DT on the basis

of 12 items, four per trait. The DD has demonstrated a replicable three-factor structure and adequate internal consistency, test-retest reliability, and validity (Jonason & Webster, 2010; Jonason & Luévano, 2013; Webster & Jonason, 2013), although some concerns have been raised about the construct validity of the psychopathy and Machiavellianism scales (Furnham et al., 2013; Jones & Paulhus, 2014; Miller et al., 2012). To date, tests for sex differences in observed DD scores typically found that men score higher than women on all scales (values of Cohen's *d* ranged from .09 to .79), although this finding seems most robust with respect to psychopathy (Table 1).

[Table 1]

Klimstra et al. (2014) found evidence of measurement invariance of the DD across sex in two samples of Dutch adolescents, by reporting that boys consistently scored higher than girls on psychopathy, whilst the evidence for sex differences in Machiavellianism and narcissism was less robust. Although this study shed some light on the measurement invariance of the DD across sex, all the data were collected from an adolescent population. The measurement invariance of DD scales in adults has yet to be examined.

This Study

The aim of this study was to assess the measurement invariance of the DD across sex in an adult community sample. The sample consisted of Italian adults, but when we started this research project no validated Italian version of the DD was available. Hence, the first step was to produce an Italian translation of the DD and evaluate its psychometric properties. In Study 1, we assessed whether the Italian DD (henceforth, DD-I) replicated the three-factor structure of the original, using three independent samples of participants and exploratory and confirmatory factor analyses. In Study 2, we investigated the test-retest reliability and temporal stability of DD-I scores. In Study 3, we assessed the convergent and discriminant validity of the DD-I. The main aim of the study was addressed in Study 4, in which we assessed the measurement invariance of the DD across sex using multi-group confirmatory factor analysis.

Study 1

In Study 1, we investigated the factor structure of the Italian translation of the DD. The translation was produced through a mixed forward- and back-translation procedure (Behling & Law, 2000). The authors and three PhD students and post-doctoral psychology researchers independently translated the original English-language DD scales into Italian. Once a consensus Italian version had been produced, an Italian-English bilingual speaker who was blind to the original version translated this preliminary Italian DD back into English. Discrepancies between the original version and the back-translation, and other issues relevant to the process of adapting the scale for the Italian context were discussed by the translators. A final Italian version was produced, taking into account the meanings of the original English items and ensuring that the content of the Italian translation was appropriate to the constructs being measured. Once the final Italian version had been agreed upon, it was administered to ten naïve individuals to check the clarity and readability of the items. All items were found to be easy to understand and score (all the characteristics were scored using a 10-point scale, with higher scores reflecting more positive evaluations; the mean scores were all higher than 9). The DD-I items are listed in the Appendix.

Method

Participants and Procedure

The DD-I was administered to three independent community samples from northern Italy. As part of their dissertation or research-training project, three undergraduate psychology students were asked to recruit participants from their relatives, friends, and acquaintances, and, when needed, they asked these participants to put them in contact with other potential participants. The groups should therefore be considered opportunity, convenience samples of individuals drawn from the community. Sample 1 comprised 102 participants (age M = 4.04 years, SD = 14.45, range 18-69, 53% women), Sample 2 comprised 128 participants (age M = 35.75 years, SD = 14.96 years, range 18-80, 57% women), Sample 3 comprised 305 participants (age M = 37.34 years, SD = 13.30 years, range 18-74, 61% women). All participants were volunteers and agreed to participate after being presented with a detailed description of the procedure, and participants were treated in

accordance with the *Ethical Principles of Psychologists and Code of Conduct* (American Psychological Association, 2010). The minimum age for participants was 18 years old and individuals who reported that they had been diagnosed with a psychiatric disorder were excluded. Participants did not receive any compensation. The DD-I was administered in a quiet room at the premises of a university psychology department.

Measures

Together with the questionnaire, participants found a cover letter that briefly explained that the aim of the study was to investigate individual differences in some personality characteristics. They were also informed that participation was anonymous and voluntary, that if they decided to participate or not to participate there would have been no loss of benefits to which they were otherwise entitled, that they could skip any question, and that they could decide to stop participating without consequences. Finally, it was made explicit that the results would be reported in aggregate form only, that participants could not be identified individually, and that if the participant decided to complete the questionnaire she/he was implicitly giving their consent to participate in the study.

Dirty Dozen. In Study 1 the only instrument administered was the DD-I. Like the original DD (Jonason & Webster, 2010), the DD-I consists of 12 items to be rated using a seven-point Likert scale ranging from 1 = "*strongly disagree*" to 7 = "*strongly agree*". More information about the scale is given in the Introduction.

Data Analysis

With the data from samples 1 and 2 we took an *exploratory*, rather than a confirmatory, approach to the assessment of the factor structure of the DD-I. Confirmatory factor analysis (CFA) requires that each indicator loads on only one factor but, as recent studies have shown (Asparouhov & Muthén, 2009), this assumption might be too restrictive for personality research, where many indicators tend to have secondary loadings significantly different from zero. The presence of secondary loadings is a critical issue, since it implies that an item has weak discriminant validity, namely, it can be considered an indicator of the secondary-loading construct(s) as well as the target

construct. In CFA, the greater the departure of the secondary loadings from zero, the greater the inflation of correlations among the factors, in order to account for non-zero secondary loadings that have been constrained to be zero. This may result in biased loadings, overestimation of factor correlations, distortion of structural relationships, and lack of fit (Asparouhov & Muthén, 2009). Jonason and Webster (2010) found some evidence of substantial (i.e., larger than |.30|) cross-loadings in the DD (see their Table 2, p. 423).

Another reason for choosing an exploratory approach was that Samples 1 and 2 did not afford sufficient statistical power to estimate all the model parameters and, particularly, their standard errors (Muthén & Muthén, 2002). We performed a Monte Carlo analysis based on the results of preliminary factor analyses (see Muthén & Muthén, 2002 for details of this procedure). Using the criteria suggested by these authors, we calculated that we would need a minimum of 300 participants per group to achieve a power of .80 (Section 1 of the Electronic Supplementary Materials [ESM]). However, according to de Winter, Doudou and Wieringa (2009), the sample sizes were adequate for maximum likelihood exploratory factor analysis (ML-EFA), as they recommend a minimum of 67 participants for assessing data with a three-factor structure, based on 12 items with factor loadings in the .60s (de Winter et al., 2009, p. 155). The analyses were performed with the *fa* function in the R package *psych* (Revelle, 2015). An oblique promax rotation was applied. We used parallel analysis (PA; Horn, 1965) and the minimum average partial correlation statistic (MAP; Velicer, 1976) to determine the optimal number of factors.

To test the similarity of the factor solutions in the two samples, and hence their replicability in independent samples, we computed congruence coefficients (CCs; Tucker, 1951). CCs are a measure of factor similarity that can be used when data do not meet the requirements for structural equation modeling (SEM) (Lorenzo-Seva & ten Berge 2006), as it was the case here. CCs can be interpreted as a standardized measure of the proportionality of elements in factor loading matrices for different samples, and they measure factor similarity independently from the mean absolute size of the loadings. CCs range from −1 to 1 and values in the range .85–.94 suggest adequate similarity,

whilst values higher than .95 suggest broadly equivalent factor loading matrices (Lorenzo-Seva & ten Berge 2006).

The data from Sample 3 were subjected to CFA as they afforded sufficient statistical power. We also tested the fit of two more parsimonious alternatives to the three-correlated-factors model, namely, a one-factor model and a three-independent-factors model. We also tested the fit of a bifactor model that outperformed the other measurement models for DD items in recent studies (Czarna et al., 2016; Jonason & Luévano, 2013). In the bifactor model items load on two types of latent factors: a single, latent, general DT factor, and three latent factors associated with specific DT traits. The general DT factor can be interpreted as a measure of the residual DT covariance after the variance attributable to the three latent trait factors has been removed. Similarly, the three latent trait factors reflect their respective measures after removing the variance attributed to the single DT factor. For model identification purposes all latent factors were left uncorrelated. Diagrams of these models are shown in Figure 1.

[Figure 1]

CFA was performed with Mplus 7 (Muthén & Muthén, 1998–2012). We used the Mplus robust maximum likelihood estimator (MLR), with standard errors and tests of fit that were robust to the non-normality of observations. The goodness-of-fit of CFA models was evaluated using the comparative fit index (CFI), the Tucker-Lewis index (TLI), and the root-mean-square error of approximation (RMSEA). We used the following criteria for model fit (Marsh, Hau, & Wen, 2004): TLI and CFI: values ≥ .90 indicate acceptable fit, values ≥ .95 indicate excellent fit; RMSEA: values ≤ .08 indicate acceptable fit, values ≤ .06 indicate excellent fit.

Results

Descriptive statistics showed that the distribution of the data was characterized by moderate positive skewness (Sample 1: median = 1.04, *M* = .95; range: .26-1.51; Sample 2: median = .67, *M* = .79, range: .03-1.81) and moderate negative kurtosis (Sample 1: median = -.19, *M* = -.10, range: -

1.17-1.46; Sample 2: median = -.45, *M* = .07, range: -.95-2.92). More details are reported in Table 2.

[Table 2]

Dimensionality analyses provided convincing evidence of the adequacy of a three-factor structure. Indeed, in both samples the scree plot began to level off after the third factor, only the first three observed eigenvalues were higher than the simulated ones, and the MAP reached its minimum at three components (Section 2 of the ESM). On the basis of these results we extracted three factors in the exploratory factor analyses (EFAs). The three-factor solution accounted for 55% of variance in DD-I scores in Sample 1 and 62% of variance in DD-I scores in Sample 2. Factor loadings and factor correlations are reported in Table 3.

[Table 3]

All items loaded substantially on the expected factor (Sample 1, median target loading: .71, range: .48-.96; Sample 2, median target loading: .79, range: .51-.94), with minimal cross-loadings (Sample 1, median cross-loading: .01, range: -.27-.33; Sample 2, median cross-loading: .04, range: -.19-.26). CCs for the three factors were all .95. Correlations between the Machiavellianism factor and the other factors were around .50, whilst the correlation between narcissism and psychopathy was somewhat lower (.40 in Sample 1 and .29 in Sample 2); both results are consistent with previous research (Jonason & Webster, 2010). The two correlation matrices had similar coefficients ($X^2(6) = 3.24$, $p = .222$), suggesting that the pattern of associations between factor scores was stable. Taken together, these results suggest that the DD-I has a robust, three-factor structure similar to that of the original DD.

The results of the CFAs are reported in Table 4 and suggested that the bifactor model had the best fit. However, estimated factor loadings for items 1, 4, and 5 were not statistically significant in this model (Table 5). The three-correlated-factors model had the next-best fit and parameter estimates of factor loadings and factor correlations were all statistically significant (Table 5). The size of the factor correlations was consistent with results in Samples 1 and 2.

[Table 4]; [Table 5]

Next, we performed an item analysis on observed scores. We computed values of Cronbach's alpha, mean inter-item correlations, corrected-item total correlations, items' squared multiple correlations, and alpha-if-item-deleted indices for all DD-I scales in all samples. Detailed results are reported in Table 6 and show that, despite the relatively small number of items, the DD-I scales have a high degree of internal consistency. Values of Cronbach's alphas were equal to or larger than .80, except in the case of the Sample 3 data for the Psychopathy scale (adequate; α = .73). Corrected item-total correlations were all well above .30, indicating that the DD-I items discriminated adequately between high and low levels of the traits.

[Table 6]

Study 2

In Study 2 we tested the temporal stability of DD-I scores in an independent sample. The DD-I was the only measure used in this study and was used exactly as described in Study 1, including the cover letter.

Method

Participants and Procedure

The DD-I was administered twice to 164 psychology undergraduate students (age *M* = 22.68 years, *SD* = 5.50, range: 19-59, 77% women) from a university in northern Italy with a three-week interval between administrations. None of the participants had taken part in Study 1. Students were informed that completion of the DD-I was not compulsory, that participation would not affect their final evaluation, and that they could withdraw from the study at any time without consequences.

Results

We used EFAs to determine whether DD-I data collected at Time 1 and Time 2 shared the same factor structure. The three-factor solution accounted for 65% and 72% of variance in DD-I scores at Times 1 and 2, respectively, and the dimensionality analyses suggested that the optimal number of factors was three (Section 3 of the ESM). The EFA results are reported in Table 7 and

suggest that the factor structure of DD-I data is stable across time. We also computed CCs for the Time 1 and Time 2 factor loading matrices: the values were .99, .98, and .98 for Machiavellianism, psychopathy, and narcissism, respectively. These results suggest that the factor congruence between Times 1 and 2 is high, and this can be interpreted as evidence that the factor structure of the DD-I shows adequate temporal stability. CCs for the congruence between factor loading matrices at Times 1 and 2, on the one hand, and those of Samples 1 and 2 in Study 1, on the other, are reported at the bottom of Table 7.

[Table 7]

All CCs were higher than .86, suggesting that the DD-I factor structure in the convenience, community samples was adequately congruent with that in the student sample. Test-retest reliability was assessed by computing intraclass correlation coefficients (ICCs). ICCs were computed as single measure using a two-way random effects model with an absolute agreement definition. The results indicated that observed DD-I scores were stable over the three-week interval (ICCs ranged from .83 to .86). Paired-sample *t*-tests also indicated that the mean scores at Times 1 and 2 were similar (Table 8). The internal consistency coefficients for all scales were always greater than .8. Taken together, the results of Study 2 indicate that DD-I scores show adequate temporal stability, at least in a student population.

[Table 8]

Study 3

The construct validity of the DD-I was assessed in Study 3, using a strategy similar to that adopted by Jonason and Webster (2010). We administered a set of questionnaires that included other instruments designed to measure the same constructs as the DD-I (convergent validity), as well as instruments designed to measure constructs traditionally considered to be part of the nomological network surrounding the DT (convergent and discriminant validity). These instruments included measures of the Big Five, aggressiveness, socio-sexual orientation, self-esteem, social desirability, and impression management.

Given that disagreeableness is considered as a common core of the DT traits, we expected the DD scale scores to be negatively associated with agreeableness, a consistent finding in earlier studies (Furnham et al., 2013; Jonason, Li & Teicher, 2010; Paulhus & Williams, 2002). Previous studies also indicated that there is a robust negative association between DT traits and conscientiousness (Furnham et al., 2013, Paulhus & Williams, 2002), which is consistent with the negative associations between DT traits (especially for psychopathy) and dutifulness and deliberation (Miller et al., 2010). The pattern of associations between DD scores and other Big Five traits is less consistent (Furnham et al., 2013; Jonason & Webster, 2010).

Paulhus and Williams (2002) suggested that the DT traits also shared an aggressive core. As aggression is positively related to psychopathy (e.g. Jones & Paulhus, 2010) and narcissism (e.g. Twenge & Campbell, 2003), and Machiavellianism has been linked to the manipulative use of aggression to feign incompetence (Jones & Paulhus, 2009), we expected all DD scores to be positively correlated with self-reported aggression. This would be in line with Jonason and Webster's (2010) finding that DD scores were positively correlated with Aggression Questionnaire (AQ; Buss & Perry, 1992) scores.

DT traits have also been associated with sociosexuality (Snyder, Simpson, & Gangestad, 1986). Sociosexuality is usually seen as a dimension spanning from *restricted* to *unrestricted*. Restricted sociosexuality is characterized by long-term mating, commitment, and closeness in a relationship prior to engaging in sex with a romantic partner. Unrestricted sociosexuality is characterized by feeling relatively comfortable engaging in sex without commitment or closeness and in short-term mating. Short-term mating has been found to be associated with DT traits, especially in men (Jonason, Li, Webster, & Schmitt, 2009; Webster & Bryan, 2007). Jonason, Li, and Buss (2010) reported that individuals with high levels of DT traits, especially psychopathy, were more likely to poach mates or be poached for mating than those with lower levels of DT traits. This behavior seems to foster reproductive success and to maintain the short-term mating style in the gene pool (e.g. Mealey et al., 1995). Consistent with these results, Jonason and Webster (2010)

found that DD scores were positively associated with short-term mating (unrestricted sociosexuality). We thus expected that higher scores on the DD-I scales would be associated with higher self-reported levels of unrestricted sociosexuality.

Previous research has not provided convincing evidence of an association between DT traits and self-esteem (Hunter, Gerbing, & Boster, 1982; Fernandez & Marshall, 2003; Raskin, Novacek, & Hogan, 1991) and a more recent study (Jonason & Webster, 2010) also showed that DD scores were not correlated with measures of self-esteem. We therefore predicted that DD-I scores would not correlate with self-esteem.

Earlier studies have reported that perfectionist self-presentation, non-disclosure of imperfections, and concealment of imperfections (i.e., positive impression management; Sherry, Hewitt, Besser, Flett, & Klein, 2006) are common in individuals with higher levels of Machiavellianism (Lopes & Fletcher, 2004) and narcissism (Rauthmann, 2011). We therefore predicted that the DD scales measuring these constructs would be positively associated with impression management. It should be noted, however, that meta-analytic studies have concluded that psychopathy is only weakly associated with impression management (e.g., Ray et al. 2013). We also predicted that moralistic bias, which is the tendency to exaggerate communion-related traits such as duty, agreeableness, and impulse control (Paulhus, 2002), would be negatively correlated with DD scores, as it is characterized by high levels of ego control, achievement via conformity, nurturance, social closeness, interpersonal sensitivity, restraint, and socialization (Paulhus & John, 1998).

Method

Participants and Procedure

Sixty-six adults from northern Italy (*M* age = 32.06 years, *SD* = 12.50, range 20-59 years; women = 67%) took part in Study 3. Participants were recruited by an undergraduate psychology student from among his/her relatives, friends, and acquaintances as part of his/her dissertation project. None of the participants had taken part in Study 1 or Study 2. Like the Study 1 sample, the

Study 3 sample was an opportunity, convenience sample drawn from the community. The inclusion and exclusion criteria and the cover letter were the same as in Study 1. All participants were tested individually and anonymously in a quiet room at the premises of a psychology department. The scales making up the battery were administered in counterbalanced fashion to control for order and sequence effects.

Measures

Italian Dirty Dozen (DD-I). As described above. Descriptive statistics and Cronbach's alpha values for the DD-I and all the other instruments used are reported in Table 9.

[Table 9]

Multidimensional Personality Profile (MPP; Caprara, Barbaranelli, De Carlo & Robusto, 2006). The Machiavellianism, social desirability, and impression management subscales of the MPP, an Italian measure of personality traits, were used in this study. The Machiavellianism subscale assesses the tendency to put one's own needs ahead of those of others and to use manipulation, deceit, and tactics (e.g., bending the rules) to achieve one's goals. The social desirability subscale is a measure of moralistic bias, which is a self-deceptive tendency to deny socially deviant impulses and behaviors, and to claim saintly attributes (Paulhus & John, 1998). The impression management subscale is a measure of egoistic bias, which is a self-deceptive tendency to exaggerate one's social and intellectual status (Paulhus & John, 1998). In all these subscales respondents are required to indicate to what extent each statement applies to them using a five-point, Likert-type scale. After recoding of reverse-scored items, item scores are summed to yield a total score for each scale, and higher scores indicate higher levels of the relevant trait. As reported in the manual (Caprara et al., 2006), the three scales had adequate discriminant validity with respect to measures of agency, self-regulation, innovation, emotional intelligence, and coping ability. Cronbach's alphas in the normative sample were .63, .69, and .66 for Machiavellianism, social desirability, and impression management, respectively.

Psychopathic deviate subscale of the Minnesota Multiphasic Personality Inventory-2 (MMPI-PD; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989; Italian version: Pancheri & Sirigatti, 1995). The psychopathic deviate subscale of the MMPI consists of 50 true-false items that are used to assess social deviation, denial of authority, and amorality. After recoding of the reverse-scored items, item scores are summed to yield a total score, and higher scores indicate higher levels of psychopathic tendencies. The Italian version of the MMPI-PD showed adequate internal consistency (ranging from .65 to .70, Pancheri & Sirigatti, 1995), and factor analyses showed that it positively loaded on a general psychotic factor along with paranoia, schizophrenia, and the faking bad (F) validity scale, and it negatively loaded on a phobic experience factor along with fears and social discomfort (Donà, Micheluzzi, & Boaretto, 2006).

Narcissistic Personality Inventory (NPI; Raskin & Hall, 1988; Italian version: Fossati & Borroni, 2008a). The NPI is a widely used measure of narcissism. It comprises 40 items each requiring the respondent to choose which of two statements is more applicable to him or her. One of each pair of statements reflects a narcissistic attitude and the other does not. The score is the number of narcissistic statements selected, and higher scores indicate higher levels of narcissism. As reported by Fossati and Borroni (2008a), the Italian NPI showed a replicable one-factor structure across four independent samples, the Cronbach's alpha of the scores ranged from .81 to .89, and the scale score showed a significant positive correlation with the score on the narcissistic dimension of the SCID-II 2.0 (First, Gibbon, Spitzer, Williams, & Benjamin, 1997).

Big Five Inventory (BFI; John, Donahue & Kentle, 1991; Italian version: Ubbiali, Chiorri, Hampton, & Donati, 2013). The BFI is a 44-item self-report measure of the Big Five (Extraversion [8 items], Agreeableness [9 items], Conscientiousness [9 items], Neuroticism [8 items] and Openness [10 items]). The items are short statements that include trait adjectives recognized as prototypical of Big Five traits and respondents are asked to indicate the extent to which they agree with each statement using a five-point, Likert-type scale. After recoding of the reverse-scored items, item scores are summed to yield a total score for each scale, and higher scores indicate higher levels

of the relevant trait. As reported by Ubbiali et al. (2013), the Italian BFI has shown adequate psychometric properties as the results of the factor analyses replicated the expected five-factor structure of the original, Cronabach's alphas of the scales ranged from .69 to .83, test-retest reliability ranged from .79 to .97, and scale scores showed the expected pattern of correlations with the corresponding domain and facet scores of the Big Five Questionnaire (Caprara, Barbaranelli, Borgogni, & Perugini, 1993).

Aggression Questionnaire (AQ; Buss & Perry, 1992; Italian version: Fossati & Borroni, 2008b). The AQ is a 29-item, self-report measure of anger and aggression. Respondents are asked to indicate how characteristic each statement is of them using a five-point, Likert-type scale. The AQ provides scores for four subscales: physical aggression (9 items), verbal aggression (5 items), anger (7 items) and hostility (8 items). After recoding of the reverse-scored items, item scores are summed to yield a total score for each subscale, and higher scores indicate higher levels of the relevant trait. As reported by Fossati and Borroni (2008b), item-level multiple-group component analysis on the Italian AQ replicated the Buss and Perry's 4-factor structure in two non-clinical and one clinical sample and Cronbach's alphas ranged from .60 to .89. Scale scores showed significant negative correlations with measures of cooperativeness and significant positive correlations with measures of insecure attachment and narcissistic and borderline personality.

Rosenberg's Self-Esteem Scale (RSES; Rosenberg, 1965; Italian version: Prezza, Trombaccia, & Armento, 1997). The RSES is a 10-item, self-report measure of global self-esteem. Respondents are asked to indicate how strongly they agree with each statement using a four-point, Likert-type scale. After recoding of the reverse-scored items, item scores are summed to yield a total score, and higher scores indicate higher levels of self-esteem. Prezza et al. (1997) reported that the Italian RSES had a single-factor structure and an adequate Cronbach's alpha (.84), and the scale score showed the expected pattern of correlations with measures of social support, depression and anxiety symptoms, femininity and masculinity, life satisfaction, and self-esteem.

Sociosexual Orientation Inventory-Revised (SOI-R; Penke & Asendorpf, 2008; Italian version available at: http://www.larspenke.eu/en/translated-soi-r.html). The SOI-R is a measure of sociosexuality. It provides a total score and scores on three subscales: Behavior (3 items, e.g. 'With how many different partners have you had sex within the past 12 months?'), Attitude (3 items, e.g. 'Sex without love is OK'), and Desire (3 items, e.g. 'How often do you experience sexual arousal when you are in contact with someone you are not in a committed romantic relationship with?'). Items are scored using a nine-point scale. Item scores are summed to yield a total score for each scale. High scores indicate an unrestricted sociosexual orientation and low scores indicate a restricted sociosexual orientation. The Italian version of this scale is available at the original authors' website, but to the best of our knowledge no formal Italian validation study has been performed. As reported in Table 9, the total and subscale scores showed adequate levels of internal consistency. Additional analyses (see Section 4 of the ESM) showed that, at least in this study, the expected three-correlated-factor structure was replicated and that the pattern of correlations of scores with scores on other measures was consistent with results from previous studies (e.g., Penke & Asendorpf, 2008).

Results

Results are reported in Table 9. As expected, the DD-I scales were significantly correlated with their respective single-scale instruments (Machiavellianism and MPP - Machiavellianism, r = .48; psychopathy and MMPI-PD, r = .41; narcissism and NPI, r = .54). However, the DT scales were also intercorrelated (e.g., Machiavellianism and psychopathy, r = .50). The three DD-I scales were also negatively correlated with agreeableness, but not with conscientiousness, albeit the effect sizes of correlations with the psychopathy and narcissism scales (correlations in the .10s-.20s) were comparable to those of Jonason and Webster (2010), thus suggesting a possible lack of statistical power. Both Machiavellianism and narcissism were positively correlated with extraversion, which is consistent with Jonason et al. (2010). The DD-I scales were also positively associated with AQ subscales and SOI-R subscales (except SOI-R Behavior), results which are consistent with the

predictions and Jonason and Webster's (2010) results. Finally, the correlations between DD-I scale scores and the measures of social desirability and impression management were also consistent with the hypotheses. Machiavellianism and narcissism scores were negatively associated with moralistic bias and positively associated with egoistic bias, whereas psychopathy was unrelated to social desirability and impression management.

Study 4

After finding evidence of the adequacy of the psychometric properties of the DD-I, in Study 4 we could address the main aim, namely, the assessment of the measurement invariance of the DD-I across sex. The DD-I was the only measure used in Study 4.

Methods

Participants and procedure

Participants were recruited in central Italy through a convenience sampling procedure in which students enrolled in an undergraduate psychology class held by one of the Authors were asked to complete the DD-I themselves and to recruit a few other participants among their pool of friends, family, and acquaintances. They were instructed to follow the ethical principles of psychologists and code of conduct and to administer the DD-I in a quiet room. They received no compensation for this activity, and accepted to do it on a voluntary basis. In order to guarantee anonymity, the information about which participants were recruited by which student could not be recorded. The materials were the same as Study 1, including the cover letter.

The total number of participants was 974 (56.9% women; *M* age = 36.45 years, *SD* = 13.21, range: 18-80). The sex groups were adequately matched for age and other background characteristics (Section 5 of the ESM). The inclusion and exclusion criteria were the same as in Study 1.

Statistical Analyses

Multi-Group Confirmatory Factor Analysis (MG-CFA) was used to assess the measurement invariance of the DD-I. First, we tested the fit of the *a priori* three-correlated-factors model to the

data for the whole sample and for the separate sex groups, and then we assessed measurement invariance across sex. As a first step, the hypothesized factor structure was estimated simultaneously in women and men (*configural invariance model*, M0). This model tested whether the same factor structure was maintained across groups. We then constrained the factor loadings to be equal (*weak invariance model*, M1) to determine whether item scores increased proportionally in women and men for a given increase in the latent factor. Comparisons of *latent* scores would be warranted only if women and men with similar levels of the construct had comparable scores on items reflecting the construct, so item intercepts (i.e., the expected item scores when the score on the latent factor is zero) were constrained to be invariant (*strong invariance model*, M2). Comparison of *manifest* scores would be warranted if the constructs had similar levels of measurement error in women and men, and so items' residual variances were also constrained to be invariant (*strict invariance model*, M3).

Since we wanted to extend previous results to a larger range of invariance models, we also tested models in which latent factor variances (M4) and covariances (M5) were constrained to be invariant. M4 implied that women and men used the same range on the factor continuum when reporting DT traits and that items had equal reliability in both sexes. M5 assumed that the correlations between factor pairs were the same in women and men (for more details see Section 6 of the ESM).

The fit of the CFA models was evaluated using the same criteria as in Study 1. Measurement invariance models were also compared using fit indices. Research has shown that model comparisons based on a chi-squared difference test are affected by the same problems as the chi-squared goodness of fit test, which led to the development of fit indices (see Marsh, Hau, Balla, & Grayson, 1998). We therefore accepted a change in CFI of less than .01 or a change in RMSEA of less than .015 (Chen, 2007) as evidence of model invariance. These are commonly accepted *ad hoc* criteria for deciding when differences in fit are large enough to warrant the rejection of a more parsimonious model in favor of a more complex model (see, e.g. Marsh et al. 2010).

Results

Descriptive statistics for the sample as a whole and for the sex groups are reported in Table 10. First, we tested the fit of a three-correlated-factors CFA model to data from the whole sample and data for the sex groups. The results showed that the hypothesized model had an adequate fit in all cases (Table 11).

[Table 10]; [Table 11]

We then assessed the measurement invariance of the DD-I across sex. Factor loadings, item intercepts, residual variances, and factor correlations for the configural invariance model (M0) are reported in Section 7 of the ESM. All parameter estimates were statistically different from zero ($p < .001$). Inspection of the fit indices for the invariance models reported in Table 11 revealed that the invariance of the DD-I measurement model across sex was fully supported for all models. As shown in the rightmost columns of Table 11, invariance models in which the mean differences in factor scores were estimated revealed that men had higher scores than women on all factors, although the effect size was small in the case of narcissism. Note that in Table 11 the standardized mean difference estimates are expressed as values of Cohen's *d*.

Discussion

In earlier research, sex differences in the DT traits in general, and as captured by the DD in particular, have been assessed on the basis of observed scores, overlooking the possibility that lack of measurement invariance in the scale may have biased the results. The only exception was a study by Klimstra et al.'s (2014) on adolescents, which found evidence for the measurement invariance of the DD across sex and reported that boys tended to have higher scores than girls on all scales, especially the psychopathy scale. The aim of this study was to replicate their results in a large adult sample using an Italian translation of the DD, and to extend them by assessing a larger range of invariance models.

Before we could assess the measurement invariance of the DD, we had to produce an Italian version of the DD. Three preliminary studies suggested that the DD-I had adequate psychometric

properties. Study 1 showed that the DD-I items shared the three-factor structure of the original and that this structure was fairly robust. Unlike other studies (e.g., Czarna et al., 2016; Jonason & Luévano, 2013), we did not find convincing support for a bifactor model, as not all the bifactor model parameters were statistically significant. Study 2 confirmed that the DD-I had adequate test-retest reliability and that scores are fairly stable over a three-week period in absence of true change. Although Study 3 provided moderately good evidence for the convergent validity of the DD-I (moderate-to-strong correlations between DD-I scale scores and alternative measures of the constructs), there was little evidence for the discriminant validity of the DD-I. DD-I trait score often correlated as strongly with measures of other constructs as with the target construct. Grounding on previous studies, it might be hypothesized that the relatively poor discriminant validity may be a general limitation of the DD (e.g. Miller et al., 2012), rather than being a problem specific to this study or to the Italian version of the DD. Indeed, it should be noted that the present results are consistent with earlier studies (e.g. Czarna et al., 2016; Jonason et al., 2010; Küfner, Dufner, & Back, 2014). Czarna et al. (2016) argued that such findings might arise because the DD, which relies on a very small number of items per trait, captures only the core aspects of the traits, whereas other DT instruments consisting of a greater number of items provide a broader coverage of the DT traits. A potentially problematic result from Study 3 is that the largest pairwise correlation between DD-I scale scores was between Machiavellianism and narcissism (r = .66). However, this seems to be an idiosyncratic result. In Table 8.1 of Section 8 of the ESM we report pairwise correlations between observed DD-I scale scores in all studies. They are in the .40-.60 range for Machiavellianism-psychopathy (M-P) and Machiavellianism-narcissism (M-N), and in the .20-.40 range for narcissism-psychopathy (N-P). These results seem to be consistent (although slightly smaller in size), with those of previous studies on the DD, in which the median correlations were: M-P = .57, M-N = .52, and N-P = .33 (Table 8.2 Section 8 of the ESM). However, a recent review on the DT (O'Boyle, Forsyth, Banks, & McDaniel, 2012) reported the following pooled pairwise correlations between DT traits: M-P = .46, M-N = .23, and N-P = .42. Furnham and colleagues

(2014) reported that, although the measurement of the Dark Triad varies by study, the "overall highest correlations appear to be between the measure of psychopathy and Machiavellianism, and the lowest between narcissism and Machiavellianism" (p. 115). While the former result is replicated in studies on the DD, the latter is not, and this seem to raise further issues about the construct validity of the specific DD scales. This peculiarity might be due to the way in which the DD items were selected. Jonason and Webster (2010) initially administered the NPI, the 31-item Self-Report Psychopathy Scale–III (SRP-III; Paulhus, Hemphill, & Hare, 2009), and the 20-item Mach IV (Christie & Geis, 1970), carried out principal components analyses, and chose from each of the three Dark Triad measures the four items with the strongest loadings on the primary factor. As pointed out by Widaman, Little, Preacher, and Sawalani (2011), this strategy has shortcomings. It bases decisions on patterns of results from a particular set of data, and since the subset of items that appears to be optimal might vary across different sets of empirical data, it can capitalize on chance results in a single sample - that in the case of Webster and Jonason (2010) comprised psychology students. Moreover, it may result in a narrowing of item content, thus restricting the breadth of the item content in the full scale. As a result, the content of Machiavellianism items in the DD (manipulation and exploitation of others) may map a portion of the content domain of the construct that overlaps with the portion of the content domain mapped by DD narcissism items (seeking admiration and prestige) more than it has been observed with the original, longer measures.

The observed bivariate associations between the DT traits and external correlates were meaningful and consistent with prior studies on the DD (Jonason & Webster, 2010). Machiavellianism, psychopathy, and narcissism were negatively related to agreeableness, which is consistent with Paulhus and Williams's (2002) suggestion that a disagreeable attitude toward others might be a common feature of all DT traits. Furthermore, Machiavellianism and narcissism, but not psychopathy, were positively associated with extraversion, mirroring findings in a sample of adolescents (Klimstra et al., 2014). It should be noted, however, that Machiavellianism is not consistently associated with extraversion (e.g. O'Boyle, Forsyth, Banks, Story, & White, 2015). We

also expected negative correlations with conscientiousness, but, although the coefficients were in the expected direction and their size was consistent with previous studies on the DD (e.g., Jonason & Webster, 2010), they were not statistically significant. Since low conscientiousness is a key characteristic of the Dark Triad person (Furnham et al., 2014), further studies are needed to investigate whether this result is due to a lack of power of this study, to the above-mentioned content limitation of the DD, or to sampling biases (see below for more details on this issue).

Aggression dimensions and unrestricted sexual orientation were also positively correlated with DT traits, confirming and extending the extant evidence of the potentially risky interpersonal consequences of sub-clinical levels of Machiavellianism, psychopathy, and narcissism. Finally, Machiavellianism and narcissism – but not psychopathy – were positively associated with impression management and negatively associated with social desirability. This may indicate that these two 'dark' traits are associated with a tendency to exaggerate personal attributes and status and a willingness to acknowledge antagonistic impulses. The lack of correlation between psychopathy scores and social desirability scales is not surprising, as it has been reported that self-report measures of psychopathy tend to be only weakly negatively correlated with measures of social desirability or impression management (e.g., Lilienfeld & Fowler, 2005).

The results of the measurement invariance analyses (Study 4) suggested that the measurement model of the DD-I and its parameters are invariant across sex. Specifically, we found that factor variances and covariances are substantially invariant between women and men. This implies that (1) women and men use the same range on the factor continuum when reporting DT traits (i.e. narcissism, psychopathy, and Machiavellianism) and that items tend to have similar reliabilities in men and women; (2) the correlation between given factor pairs is statistically equivalent in women and men.

These results imply that differences in mean scores accurately reflect differences in the constructs as operationalized by the DD. Consistent with previous studies (Furnham & Trickey, 2011), we found that men scored higher than women on all DD-I scales, and that the effect sizes for

the sex differences were higher for Machiavellianism and psychopathy (.40s) than for narcissism (.15s). This result may be attributable to the multidimensional nature of narcissism, as recent research has suggested that there are two phenotypic expressions of narcissism, characterized by grandiose and vulnerable features (Cain et al., 2008). Men usually score higher than women on the grandiose dimension of narcissism, whereas smaller or no sex differences are typical for features of vulnerable narcissism (Grijalva et al., 2015). The DD narcissism scale appears to capture both facets of narcissism (Maples, Lamkin, & Miller, 2014), so the smaller sex differences on the narcissism scale (compared with the other two scales) may reflect the contributions of both grandiose and vulnerable features (which are likely to differ across sex) to DD narcissism scores. As argued in the Introduction, sex differences in the DT traits can be explained by a biosocial approach to social role theory. If individuals are socially penalized for deviating from sex role norms women may experience societal pressure to exhibit communal behaviors and face disapproval if they display agentic behaviors and hence women may be less likely to possess - or at least to report - DT traits. Given current sex role norms it is not surprising that "the entire construct of Machiavellianism [is considered] more appropriate for men than for women" (Wilson, Near & Miller, 1996, p. 293).

Limitations

Some limitations warrant mention. First, the samples used in Studies 1, 3, and 4 were opportunity, convenience samples drawn from the community, whereas the Study 2 sample consisted entirely of undergraduate psychology students. This is should be carefully considered when interpreting the results of these studies, since participants were likely to possess certain similar characteristics that were unrepresentative of the Italian population due to environmental, cultural, or socio-economic factors. Hence, the results cannot be generalized to other populations in which a measure of DT traits can be used (e.g., forensic or correctional).

Second, socially desirable responding tends to be positively related to age and negatively related to self-reports of undesirable characteristics. We relied on self-report instruments, which

may have elicited manipulative behavior from people high in DT traits. Third, the brevity of the DD-I may compromise its ability to capture the full breadth of the DT (e.g., Miller et al., 2012) and hence its ability to disentangle specific facets of the personality styles which make up the DT construct. It should be noted, however, that the present results are consistent with those obtained using longer measures.

Conclusions

These limitations notwithstanding, this study showed that all the DD-I factors capture the same construct in adult women and men, suggesting that sex differences in scale scores reported in earlier studies might not be artifacts of measurement error. Machiavellianism, psychopathy, and narcissism seem to manifest in the same way in women and men, although men report higher levels of all three traits.

Compliance with Ethical Standards

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Ethical approval: All procedures performed in the studies were in accordance with the Ethical principles of psychologists and code of conduct (American Psychological Association, 2010), with the ethical standards of the institutional and/or national research committee, and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

References

Aghababaei, N., Mohammadtabar, S., & Saffarinia, M. (2014). Dirty Dozen vs. the H factor: Comparison of the Dark Triad and Honesty-Humility in prosociality, religiosity, and happiness. *Personality and Individual Differences, 67,* 6–1. doi:1.1016/j.paid.2014.03.026

American Psychological Association. (2010). *Ethical principles of psychologists and code of conduct*. Retrieved from http://apa.org/ethics/code/index.aspx

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling, 16*(3), 397–438. http://doi.org/1.1080/10705510903008204

Behling, O., & Law, K. S. (2000). *Translating questionnaires and other research instruments: Problems and solutions*. Thousand Oaks, CA: Sage.

Buss, A. H., & Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology*, *63*(3), 452–459. http://doi.org/1.1037/0022-3514.63.3.452

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *MMPI-2: Manual for administration and scoring*. Minneapolis, MN: University of Minnesota Press.

Cain, N. M., Pincus, A. L., & Ansell, E. B. (2008). Narcissism at the crossroads: phenotypic description of pathological narcissism across clinical theory, social/personality psychology, and psychiatric diagnosis. *Clinical Psychology Review*, *28*(4), 638–656. http://doi.org/1.1016/j.cpr.2007.09.006

Cale, E. M., & Lilienfeld, S. O. (2002). Sex differences in psychopathy and antisocial personality disorder: A review and integration. *Clinical Psychology Review*, *22*(8), 1179–1207. http://doi.org/1.1016/S0272-7358(01)00125-8

Caprara, G. V., Barbaranelli, C., Borgogni, L., & Perugini, M. (1993). The "Big Five Questionnaire": A new questionnaire to assess the five factor model. *Personality and Individual Differences*, *15*(3), 281–288. http://doi.org/1.1016/0191-8869(93)90218-R

Caprara, G. V., Barbaranelli, C., De Carlo, N. A., & Robusto, E. (Eds.). (2006). *Multidimensional Personality Profile*. Milano: Franco Angeli.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, *14*(3), 464–504. http://doi.org/1.1080/10705510701301834

Christie, R., & Geis, F. L. (1970). *Studies in Machiavellianism*. New York, NY: Academic Press.

Czarna, A. Z., Jonason, P. K., Dufner, M., & Kossowska, M. (2016). The Dirty Dozen Scale: Validation of a Polish version and extension of the nomological net. *Frontiers in Psychology, 7*:445. http://doi.org/ 1.3389/fpsyg.2016.00445

De Winter, J. C. F., Dodou, D., & Wieringa, P. A. (2009). Exploratory factor analysis with small sample sizes. *Multivariate Behavioral Research*, *44*(2), 147–181. http://doi.org/1.1080/00273170902794206

Donà, G., Micheluzzi, F., & Boaretto, M. (2006). La dimensione strutturale del test MMPI-2: monofattoriale o multifattoriale? Un'analisi quantitativa e qualitativa dei fattori significativi [MMPI-2 structural dimension: monofactorial or multifactorial? Quantitative and qualitative analysis of significant factors]. *Giornale Italiano di Psicopatologia, 12,* 293-302.

Fernandez, Y. M., & Marshall, W. L. (2003). Victim empathy, social self-esteem, and psychopathy in rapists. *Sexual Abuse: Journal of Research and Treatment*, *15*(1), 11–26. http://doi.org/1.1023/A:1020611606754

First, M. B., Gibbon, M., Spitzer, R. L., Williams, J. B. W., & Benjamin, L. S. (1997). *Structured Clinical Interview for DSM-IV Axis II Personality Disorders, (SCID-II)*. Washington, DC: American Psychiatric Press, Inc.

Fossati, A., & Borroni, S. (2008a). Versione italiana del Narcissistic Personality Inventory [Italian version of the Narcissistic Personality Inventory]. In C. Maffei (Ed.), *Borderline* (pp. 387–413). Milano: Raffaello Cortina.

Fossati, A., & Borroni, S. (2008b). Versione italiana dell'Aggression Questionnaire [Italian version of the Aggression Questionnaire]. In C. Maffei (Ed.), *Borderline* (pp. 279–307). Milano: Raffaello Cortina.

Furnham, A., & Trickey, G. (2011). Sex differences in the dark side traits. *Personality and Individual Differences*, *50*(4), 517–522. http://doi.org/1.1016/j.paid.201.11.021

Furnham, A., Richards, S. C., & Paulhus, D. L. (2013). The Dark Triad of personality: A 10-year review. *Social and Personality Psychology Compass*, *7*(3), 199–216. http://doi.org/1.1111/spc3.12018

Furnham, A., Richards, S., Rangel, L., & Jones, D. N. (2014). Measuring malevolence: Quantitative issues surrounding the Dark Triad of personality. *Personality and Individual Differences*, *67*, 114–121. http://doi.org/1.1016/j.paid.2014.02.001

Grijalva, E., Newman, D. A., Tay, L., Donnellan, M. B., Harms, P. D., Robins, R. W., & Yan, T. (2015). Sex differences in narcissism: A meta-analytic review. *Psychological Bulletin*, *141*(2), 261–31. http://doi.org/1.1037/a0038231

Hare, R. D., & Neumann, C. S. (2008). Psychopathy as a clinical and empirical construct. *Annual Review of Clinical Psychology*, *4*, 217–246. http://doi.org/1.1146/annurev.clinpsy.3.022806.091452

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179–185. http://doi.org/1.1007/BF02289447

Hunter, J. E., Gerbing, D. W., & Boster, F. J. (1982). Machiavellian beliefs and personality: Construct invalidity of the Machiavellianism dimension. *Journal of Personality and Social Psychology*, *43*(6), 1293–1305. http://doi.org/1.1037/0022-3514.43.6.1293

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five Inventory-Versions 4a and 54*. Berkeley, CA: University of California,Berkeley, Institute of Personality and Social Research.

Jonason, P. K., & Krause, L. (2013). The emotional deficits associated with the Dark Triad traits: Cognitive empathy, affective empathy, and alexithymia. *Personality and Individual Differences*, *55*, 532–537. http://dx.doi.org/1.1016/j.paid.2013.04.027

Jonason, P. K., & Luévano, V. X. (2013). Walking the thin line between efficiency and accuracy: Validity and structural properties of the Dirty Dozen. *Personality and Individual Differences*, *55*(1), 76–81. http://doi.org/1.1016/j.paid.2013.02.010

Jonason, P. K., & Webster, G. D. (2010). The Dirty Dozen: A concise measure of the dark triad. *Psychological Assessment*, *22*(2), 420–432. http://doi.org/1.1037/a0019265

Jonason, P. K., Li, N. P., & Buss, D. M. (2010). The costs and benefits of the Dark Triad: Implications for mate poaching and mate retention tactics. *Personality and Individual Differences*, *48*(4), 373–378. http://doi.org/1.1016/j.paid.2009.11.003

Jonason, P. K., Li, N. P., & Teicher, E. A. (2010). Who is James Bond?: The dark triad as an agentic social style. *Individual Differences Research*, *8*(2), 111–12.

Jonason, P. K., Li, N. P., Webster, G. D., & Schmitt, D. P. (2009). The Dark Triad: Facilitating a short-term mating strategy in men. *European Journal of Personality*, *23*(1), 5–18. http://doi.org/1.1002/per.698

Jones, D. N., & Paulhus, D. L. (2009). Machiavellianism. In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (pp. 93–108). New York: Guilford Press.

Jones, D. N., & Paulhus, D. L. (2010). Differentiating the Dark Triad within the interpersonal circumplex. In L. M. Horowitz & S. N. Strack (Eds.), *Handbook of interpersonal theory and research* (pp. 249–267). New York: Guilford Press.

Jones, D. N., & Paulhus, D. L. (2014). Introducing the Short Dark Triad (SD3): A brief measure of dark personality traits. *Assessment*, *21*(1), 28–41. http://doi.org/1.1177/1073191113514105

Klimstra, T. A., Sijtsema, J. J., Henrichs, J., & Cima, M. (2014). The Dark Triad of personality in adolescence: Psychometric properties of a concise measure and associations with adolescent adjustment from a multi-informant perspective. *Journal of Research in Personality*, *53*, 84–92. http://doi.org/1.1016/j.jrp.2014.09.001

Küfner, C. P., Dufner, M., and Back, M. D. (2014). Das Dreckige Dutzend und die Niederträchtigen Neun – Kurzskalen zur Erfassung von Narzissmus, Machiavellismus und Psychopathie. *Diagnostica, 61*, 79–91. http://doi.org/ 1.1026/0012-1924/a000124

Lilienfeld, S. O., & Fowler, K. A. (2005). The self-report assessment of psychopathy. Problems, pitfalls, and promises. In C. J. Patrick (Ed.), *Handbook of psychopathy* (pp. 107–132). New York: The Guilford Press.

Lopes, J., & Fletcher, C. (2004). Fairness of impression management in employment interviews: A cross-country study of the role of equity and Machiavellianism. *Social Behavior and Personality*, *32*(8), 747–768. http://doi.org/1.2224/sbp.2004.32.8.747

Lorenzo-Seva, U., & ten Berge, J. M. F. (2006). Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology*, *2*(2), 57–64. http://doi.org/1.1027/1614-2241.2.2.57

Maples, J. L., Lamkin, J., & Miller, J. D. (2014). A test of two brief measures of the Dark Triad: The Dirty Dozen and Short Dark Triad. *Psychological Assessment*, *26*(1), 326–331. http://doi.org/1.1037/a0035084

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, *11*(3), 320–341. http://doi.org/1.1207/s15328007sem1103_2

Marsh, H. W., Hau, K.-T., Balla, J. R., & Grayson, D. (1998). Is more ever too much? The number of indicators per factor in confirmatory factor analysis. *Multivariate Behavioral Research*, *33*(2), 181–22. http://doi.org/1.1207/s15327906mbr3302_1

Marsh, H. W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A. J. S., Trautwein, U., & Nagengast, B. (2010). A new look at the big five factor structure through exploratory structural equation modeling. *Psychological Assessment*, *22*(3), 471–91. http://doi.org/1.1037/a0019227

Mealey, L., Aird, R. B., Yamamoto, T., Alexander, R. D., Allen, H., Lindner, L., … Murphy, D. L. (1995). The sociobiology of sociopathy: An integrated evolutionary model. *Behavioral and Brain Sciences*, *18*(3), 523. http://doi.org/1.1017/S0140525X00039595

Miller, J. D., Dir, A., Gentile, B., Wilson, L., Pryor, L. R., & Campbell, W. K. (2010). Searching for a vulnerable Dark Triad: Comparing factor 2 psychopathy, vulnerable narcissism, and borderline personality disorder. *Journal of Personality*, *78*(5), 1529–1564. http://doi.org/1.1111/j.1467-6494.201.0066.x

Miller, J. D., Few, L. R., Seibert, L. A., Watts, A., Zeichner, A., & Lynam, D. R. (2012). An examination of the Dirty Dozen Measure of psychopathy: A cautionary tale about the costs of brief measures. *Psychological Assessment*, *24*(4), 1048–1053. http://doi.org/1.1037/a0028583.

Millsap, R. E. (2011). *Statistical approaches to measurement invariance*. New York: Taylor & Francis Ltd.

Morf, C. C., & Rhodewalt, F. (2001). Expanding the dynamic self-regulatory processing model of narcissism: Research directions for the future. *Psychological Inquiry*, *12*(4), 243–251. http://doi.org/1.1207/S15327965PLI1204_3

Muris, P., Meesters, C., & Timmermans, A. (2013). Some youths have a gloomy side: Correlates of the dark triad personality traits in non-clinical adolescents. *Child Psychiatry & Human Development, 44*(5), 658–665. doi:1.1007/s10578-013-0359-9

Muthén, B., & Muthén, L. (1998-2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Muthén, L., & Muthén, B. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling*, *9*(4), 599–62. http://doi.org/1.1207/S15328007SEM0904_8

O'Boyle, E. H, J., Forsyth, D. R., Banks, G. C., & McDaniel, M. A. (2012). A meta-analysis of the dark triad and work behavior: A social exchange perspective. *Journal of Applied Psychology, 97*(3), 557–579. http://doi.org/1.1037/a0025679

O'Boyle, E. H., Forsyth, D. R., Banks, G. C., Story, P. A., & White, C. D. (2015). A meta-analytic test of redundancy and relative importance of the dark triad and five-factor model of personality. *Journal of Personality, 83*(6), 644–664. http://doi.org/1.1111/jopy.12126

Pancheri, P., & Sirigatti, S. (Eds.). (1995). *MMPI-2 - Minnesota Multiphasic Personality Inventory - 2. Manuale*. Firenze, Italy: Giunti O.S.

Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In D. N. Jackson & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement*. Mahwah, NJ: Lawrence Erlbaum Associates.

Paulhus, D. L., & John, O. P. (1998). Egoistic and moralistic biases in self-perception: The interplay of self-deceptive styles with basic traits and motives. *Journal of Personality, 66*(6), 1025. http://doi.org/1.1111/1467-6494.00041

Paulhus, D. L., & Williams, K. M. (2002). The Dark Triad of personality: Narcissism, Machiavellianism, and psychopathy. *Journal of Research in Personality, 36*(6), 556–563. http://doi.org/1.1016/S0092-6566(02)00505-6

Paulhus, D. L., Hemphill, J. F., & Hare, R. D. (2009). *Manual for the Self-Report Psychopathy scale (SRP-III)*. Toronto, Ontario, Canada: Mulit-Health Systems.

Penke, L., & Asendorpf, J. B. (2008). Beyond global sociosexual orientations: a more differentiated look at sociosexuality and its effects on courtship and romantic relationships. *Journal of Personality and Social Psychology, 95*(5), 1113–1135. http://doi.org/1.1037/0022-3514.95.5.1113

Prezza, M., Trombaccia, F. R., & Armento, L. (1997). La scala dell'autostima di Rosenberg: traduzione e validazione italiana [The Rosenberg Self-Esteem Scale: Italian translation and validation]. *Bollettino Di Psicologia Applicata, 223*, 35–44.

Raskin, R. N., & Hall, C. S. (1979). A narcissistic personality inventory. *Psychological Reports, 45*(2), 59. http://doi.org/1.2466/pr.1979.45.2.590

Raskin, R., Novacek, J., & Hogan, R. (1991). Narcissism, self-esteem, and defensive self-enhancement. *Journal of Personality, 59*(March 1991), 19–38. http://doi.org/1.1111/j.1467-6494.1991.tb00766.x

Rauthmann, J. F. (2011). Acquisitive or protective self-presentation of dark personalities? Associations among the Dark Triad and self-monitoring. *Personality and Individual Differences, 51*(4), 502–508. http://doi.org/1.1016/j.paid.2011.05.008

Ray, J. V., Hall, J., Rivera-Hudson, N., Poythress, N. G., Lilienfeld, S. O., & Morano, M. (2013). The relation between self-reported psychopathic traits and distorted response styles: A meta-analytic review. *Personality Disorders: Theory, Research, and Treatment*, *4*(1), 1–14. http://doi.org/1.1037/a0026482

Revelle, W. (2015). *psych: Procedures for Personality and Psychological Research*. Northwestern University, Evanston, Illinois, USA, http://CRAN.R-project.org/package=psych Version = 1.5.4.

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.

Rudman, L. A., Moss-Racusin, C. A., Phelan, J. E., & Nauts, S. (2012). Status incongruity and backlash effects: Defending the gender hierarchy motivates prejudice against female leaders. *Journal of Experimental Social Psychology*, *48*(1), 165–179. http://doi.org/1.1016/j.jesp.2011.1.008

Sherry, S. B., Hewitt, P. L., Besser, A., Flett, G. L., & Klein, C. (2006). Machiavellianism, trait perfectionism, and perfectionistic self-presentation. *Personality and Individual Differences*, *40*(4), 829–839. http://doi.org/1.1016/j.paid.2005.09.010

Snyder, M., Simpson, J. A., & Gangestad, S. (1986). Personality and sexual relations. *Journal of Personality and Social Psychology*, *51*(1), 181–19. http://doi.org/1.1037/0022-3514.51.1.181

Tucker, L. R. (1951). *A method for synthesis of factor analysis studies (Personnel Research Section. Report No. 984)*. Washington, DC.

Twenge, J. M., & Campbell, W. K. (2003). "Isn't it fun to get the respect that we're going to deserve?" Narcissism, social rejection, and aggression. *Personality and Social Psychology Bulletin*, *29*(2), 261–272. http://doi.org/1.1177/0146167202239051

Ubbiali, A., Chiorri, C., Hampton, P., & Donati, D. (2013). Psychometric properties of the Italian adaptation of the Big Five Inventory (BFI). *Bollettino di Psicologia Applicata*, *266*, 37–48.

Velicer, W. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, *41*(3), 321–327. http://doi.org/1.1007/BF02293557

Velotti, P., Elison, J., & Garofalo, C. (2014). Shame and aggression: Different trajectories and implications. *Aggression and Violent Behavior*, *19*(4), 454–461. http://doi.org/1.1016/j.avb.2014.04.011

Webster, G. D., & Bryan, A. (2007). Sociosexual attitudes and behaviors: Why two factors are better than one. *Journal of Research in Personality*, *41*, 917–922. http://doi.org/1.1016/j.jrp.2006.08.007

Webster, G. D., & Jonason, P. K. (2013). Putting the "IRT" in "Dirty": Item response theory analyses of the Dark Triad Dirty Dozen-An efficient measure of narcissism, psychopathy, and Machiavellianism. *Personality and Individual Differences*, *54*(2), 302–306. http://doi.org/1.1016/j.paid.2012.08.027

Widaman, K. F., Little, T. D., Preacher, K. J., & Sawalani, G. M. (2011). On creating and using short forms of scales in secondary research. In K. H. Trzesniewski, M. B. Donnellan, & R. E. Lucas (Eds.), *Secondary data analysis: An introduction for psychologists.* (pp. 39–61). Washington: American Psychological Association. http://doi.org/1.1037/12350-003

Wiggins, J. S., & Pincus, A. L. (1989). Conceptions of personality disorders and dimensions of personality. *Psychological Assessment*, *1*(4), 305–316. http://doi.org/1.1037//1040-359.1.4.305

Wilson, D. S., Near, D., & Miller, R. R. (1996). Machiavellianism: a synthesis of the evolutionary and psychological literatures. *Psychological Bulletin*, *119*(2), 285–299. http://doi.org/1.1037/0033-2909.119.2.285

Wink, P. (1991). Two faces of narcissism. *Journal of Personality and Social Psychology*, *61*(4), 590–597. http://doi.org/1.1037/0022-3514.61.4.590

Wood, W., & Eagly, A. H. (2012). Biosocial construction of sex differences and similarities in behavior. In J. M. Olson & P. Zanna (Eds.), *Advances in experimental social psychology (Vol.46)* (pp. 55–123). Oxford, UK: Elsevier.

Figure Captions

Figure 1 Confirmatory factor analysis models in Study 1

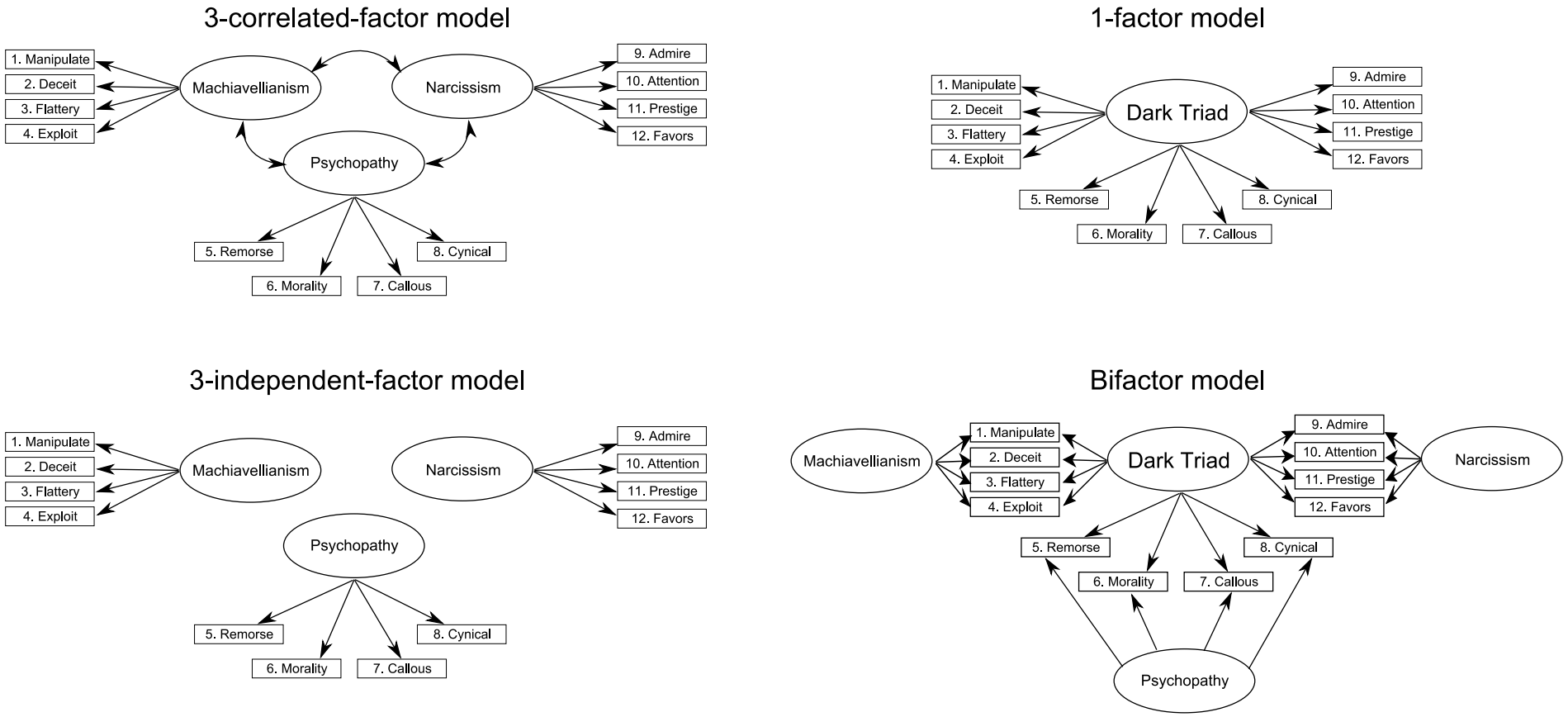

3-correlated-factor model
1. Manipulate
2. Deceit
3. Flattery
4. Exploit
Machiavellianism
Narcissism
9. Admire
10. Attention
11. Prestige
12. Favors
Psychopathy
5. Remorse
6. Morality
7. Callous
8. Cynical
1-factor model
Dark Triad
3-independent-factor model
Bifactor model

Figure 1

Table 1 Review of sex differences (Cohen's *d*) in scores on the Dirty Dozen (studies are in chronological order)

Source	Participants	Ma	P	N
Jonason & Webster (2010)				
Study 1	273 psychology students (90 men, 183 women) aged 18-47 years (M=2.08, SD=3.79)	.49	.40	.62
Study 2	246 psychology students (101 men, 145 women) aged 18-42 years (M=2.69, SD=3.76)	.21	.42	.62
Study 3	96 undergraduate students (36 men, 60 women), aged 18-25 years (M=2.44, SD=1.43)	.79	.35	.34
Study 4	470 psychology students (157 men, 312 women) aged 17-26+ years (mode=18, Median=19, M=19.00, SD=1.30)	.05	.46	.09
Jonason & Krause (2013)	320 online participants (78 men, 242 women), aged 17-56 years (M=24.24, SD= 7.33)	.75	.49	.51
Muris, Meesters & Timmermans (2013)	117 adolescents (51 men, 66 women), aged 12-18 years (M=13.90, SD=.96)	.45	.56	.01
Webster & Jonason (2013)	544 undergraduate students (169 men, 375 women), aged 17-50 (M=2.25, SD=4.70)	.31	.41	.40
Aghababaei, Mohammadtabar, & Saffarinia (2014)	223 Iranian employees (90 men, 133 women), aged 18-57 (M=31.24, SD=8.94).	.42	.42	.18
Jonason, Baughman, Carter & Parker (2015)	1,389 undergraduate students (458 men, 931 women), aged 18-50 years (M=18.88, SD=2.15)	.22	.52	.27
Czarna, Jonason, Dufner, & Kossowska (2016)				
Study 1	304 undergraduate students (111 men, 193 women) aged 18-54 years (M=22.24, SD=4.69)	.44	.54	.02
Study 2	136 undergraduate students (53 men, 83 women), aged 18-48 years M=24.40, SD=6.60)	.45	.36	.01

Note: Ma = Machiavellianism; P = Psychopathy; N = Narcissism; M = mean; SD = standard deviation; Positive *d*s indicate higher scores in males.

Table 2 Item descriptive statistics for the Italian Dirty Dozen Items in all samples of Study 1

Item	Min	Max	M	SD	SK	KU
Sample 1 (*n*=102)						
DD01	1	7	2.52	1.88	1.00	-.21
DD02	1	7	2.34	1.72	1.14	.26
DD03	1	7	2.25	1.78	1.33	.59
DD04	1	7	1.93	1.42	1.51	1.46
DD05	1	7	2.50	2.00	1.07	-.19
DD06	1	7	2.10	1.63	1.47	1.20
DD07	1	7	2.38	1.72	1.14	.35
DD08	1	7	2.59	1.88	.97	-.19
DD09	1	7	3.41	2.09	.32	-1.17
DD10	1	7	3.49	2.02	.26	-1.14
DD11	1	7	2.99	1.98	.53	-1.08
DD12	1	7	3.05	2.12	.62	-1.03
Sample 2 (*n*=128)						
DD01	1	7	2.27	1.49	1.12	.43
DD02	1	7	2.23	1.46	1.37	1.42
DD03	1	7	2.21	1.43	1.17	.74
DD04	1	7	1.86	1.30	1.81	2.92
DD05	1	7	2.67	1.53	.56	-.45
DD06	1	7	2.30	1.54	1.03	.18
DD07	1	7	2.41	1.53	.77	-.44
DD08	1	7	2.66	1.56	.56	-.65
DD09	1	7	3.72	1.80	.03	-.95
DD10	1	7	3.59	1.72	.09	-.94
DD11	1	7	3.17	1.68	.38	-.72
DD12	1	7	3.07	1.80	.54	-.72
Sample 3 (*n*=305)						
DD01	1	7	2.14	1.76	1.36	.53
DD02	1	7	1.97	1.62	1.60	1.45
DD03	1	7	2.05	1.62	1.44	1.02
DD04	1	7	1.66	1.36	2.12	3.60
DD05	1	7	2.40	1.96	1.18	.02
DD06	1	7	1.95	1.74	1.85	2.21
DD07	1	7	2.09	1.67	1.41	.80
DD08	1	7	2.18	1.83	1.44	.83
DD09	1	7	3.32	2.07	.37	-1.24
DD10	1	7	3.17	1.91	.38	-1.10
DD11	1	7	2.67	1.91	.84	-.57
DD12	1	7	2.57	1.88	.88	-.53

Note: Min = minimum; Max = maximum; M = mean; SD =standard deviation; SK = Skewness; KU = Kurtosis

Table 3 Results (factor loadings and factor correlations) of exploratory factor analyses on the Italian Dirty Dozen in samples 1 and 2 in Study 1

	Sample 1 (*n* = 102)			Sample 2 (*n* = 128)		
Item	Ma	P	N	Ma	P	N
DD01	.85	.01	-.09	.62	.07	.14
DD02	.68	.03	.11	.89	.04	-.19
DD03	.87	-.27	-.04	.81	-.13	.05
DD04	.77	.09	.02	.79	-.01	.09
DD05	-.19	.96	-.09	-.17	.70	.04
DD06	.07	.63	.00	-.05	.94	-.09
DD07	.04	.69	.00	.11	.71	.05
DD08	.33	.48	.00	.26	.51	.03
DD09	-.04	-.15	.86	.01	.00	.78
DD10	.07	.00	.65	.04	-.11	.84
DD11	.06	.06	.54	-.15	.04	.86
DD12	-.11	.05	.73	.07	.00	.75
r with P	.54			.49		
r with N	.53	.40		.56	.29	

Note: Ma = Machiavellianism; P = Psychopathy; N = Narcissism; *r* = Pearson's correlation

Table 4 Goodness-of-fit statistics of confirmatory factor analyses performed on data from Sample 3 (n = 305) in Study 1

Model and description	χ^2	df	CFI	TLI	RMSEA
1-factor model	37.77	54	.674	.601	.139
3-independent-factor model	225.88	54	.823	.784	.102
3-correlated-factor model	103.64	51	.946	.930	.058
Bifactor model	65.43	42	.976	.962	.043

Note. χ^2 = Chi-square; df = degrees of freedom; CFI = comparative fit index; TLI = Tucker–Lewis index; RMSEA = root mean square error of approximation

Table 5 Parameter estimates in the 3-correlated-factor and bifactor Confirmatory Factor Analysis (CFA) models (*n* = 305), Study 1

	Three correlated factors				Bifactor model				
	Ma	P	N	RV	G	Ma	P	N	RV
DD01	.69***	.00	.00	.52***	.67***	.21	.00	.00	.51***
DD02	.83***	.00	.00	.31***	.67***	.54***	.00	.00	.26*
DD03	.79***	.00	.00	.38***	.61***	.53**	.00	.00	.34**
DD04	.77***	.00	.00	.41***	.78***	.18	.00	.00	.36***
DD05	.00	.42***	.00	.83***	.43***	.00	.17	.00	.79***
DD06	.00	.51***	.00	.74***	.38***	.00	.34***	.00	.74***
DD07	.00	.84***	.00	.29***	.50***	.00	.64***	.00	.34***
DD08	.00	.81***	.00	.34***	.42***	.00	.77***	.00	.24
DD09	.00	.00	.73***	.47***	.35***	.00	.00	.76***	.30**
DD10	.00	.00	.75***	.44***	.50***	.00	.00	.55***	.45***
DD11	.00	.00	.66***	.57***	.47***	.00	.00	.45***	.58***
DD12	.00	.00	.73***	.47***	.57***	.00	.00	.42***	.50***
r with P	.53***				.00				
r with N	.60***	.39***			.00	.00			

Note: Ma = Machiavellianism; P = Psychopathy; N = Narcissism; RV = Residual Variance; G = General Dark Triad Factor; *r* = Pearson's correlation; * = $p < .05$; **: $p < .01$; ***: $p < .001$;

Table 6 Results of item analyses on the Italian Dirty Dozen in Study 1

Statistic	Sample 1 (*n*=102)	Sample 2 (*n*=128)	Sample 3 (*n*=305)
Mean ± SD (range)			
Ma	9.04±5.65 (4-25)	8.58±4.79 (4-28)	7.83±5.28 (4-28)
P	9.57±5.74 (4-25)	1.05±4.94 (4-24)	8.62±5.38 (4-26)
N	12.94±6.38 (4-27)	13.55±5.97 (4-28)	11.73±6.19 (4-27)
α			
Ma	.85	.86	.84
P	.80	.82	.73
N	.78	.88	.81
Mr_{ii} (range)			
Ma	.59 (.53-.67)	.62 (.53-.69)	.59 (.50-.70)
P	.50 (.36-.61)	.53 (.32-.66)	.42 (.29-.70)
N	.47 (.41-.59)	.64 (.60-.68)	.51 (.42-.61)
Mr_{it} (range)			
Ma	.69 (.64-.74)	.71 (.67-.78)	.69 (.63-.74)
P	.62 (.57-.66)	.64 (.53-.75)	.53 (.40-.62)
N	.59 (.52-.65)	.73 (.72-.76)	.63 (.58-.66)
M_{SMC} (range)			
Ma	.50 (.41-.57)	.53 (.50-.62)	.50 (.41-.57)
P	.43 (.39-.48)	.47 (.37-.58)	.36 (.18-.52)
N	.37 (.28-.46)	.54 (.52-.58)	.41 (.36-.46)
α *w/o* (highest)			
Ma	.83	.84	.84
P	.77	.81	.72
N	.76	.85	.78
Scale score *r*s			
Ma with P	.46***	.45***	.46***
Ma with N	.42***	.50***	.50***
N with P	.31**	.26**	.33***

Note: Ma = Machiavellianism; P = Psychopathy; N = Narcissism; SD = standard deviation; α = Cronbach's alpha; Mr_{ii} = Mean inter-item correlation; Mr_{it} = Mean corrected item-total correlation; M_{SMC} = Mean squared multiple correlation; α *w/o* = alpha-if-item-deleted index; *r*: Pearson correlation; * = $p < .05$; ** = $p < .01$; *** = $p < .001$.

Table 7 Factor loadings and factor correlations from exploratory factor analysis performed on data from Study 2.

	Time 1			Time 2		
Item	Ma	P	N	Ma	P	N
DD01	.66	-.01	.24	.78	.04	.12
DD02	.75	.03	-.05	.78	.03	-.01
DD03	.83	-.06	-.14	.90	-.10	-.10
DD04	.45	.38	.12	.57	.36	.06
DD05	.00	.63	.02	.07	.77	-.14
DD06	.15	.41	.06	.05	.64	.00
DD07	-.18	1.03	-.02	-.14	.96	.05
DD08	.10	.70	-.11	.01	.75	.04
DD09	-.11	-.06	.81	-.05	-.06	.85
DD10	.01	-.08	.74	.08	-.09	.75
DD11	.02	.13	.60	-.02	.09	.70
DD12	-.02	.02	.73	-.03	.04	.76
r with P	.61			.66		
r with N	.61	.35		.50	.30	
ϕ with Sample 1	.95	.89	.96	.95	.94	.98
ϕ with Sample 2	.93	.87	.97	.94	.93	.98

Note: Ma = Machiavellianism; P = Psychopathy; N = Narcissism; *r* = Pearson's correlation; ϕ = congruence coefficient

Table 8 Cronbach's alpha, descriptive statistics and intraclass correlation coefficients for Italian Dirty Dozen observed scale scores in Study 2 (n = 164)

	Time 1				Time 2						
Scale	M	SD	Range	α	M	SD	Range	α	t(163)	p	ICC
Ma	1.10	5.03	4-27	.84	1.40	5.39	4-28	.88	-1.26	.209	.83
P	9.51	5.11	4-28	.81	9.52	5.35	4-28	.83	-.06	.955	.86
N	14.36	5.45	4-28	.80	14.20	5.57	4-28	.83	.63	.533	.83

Note: M = Mean; SD = Standard deviation; α = Cronbach's alpha; Ma = Machiavellianism; P = Psychopathy; N = Narcissism; p: p-value of the paired-sample t-test; ICC = intraclass correlation coefficient

Table 9 Descriptive statistics and correlations of the Italian Dirty Dozen scale scores with scores on measures of Machiavellianism, psychopathy, narcissism, personality, aggression, self-esteem, social desirability, impression management, and sociosexual orientation (Study 3)

Variable	M	P	N	α	M	SD	Range
Ma	1.00			.80	1.27	5.20	4-25
P	.50***	1.00		.85	1.33	5.13	4-22
N	.66***	.24	1.00	.89	15.74	7.61	4-28
Convergent validity							
MPP - Machiavellianism	.48***	.33**	.28*	.75	2.02	4.15	10-28
MMPI-PD Total Score	.25*	.41**	.15	.62	2.92	4.58	12-30
NPI	.44***	.37**	.54***	.76	11.14	6.17	0-32
Discriminant validity							
BFI - Extraversion	.28*	.17	.28*	.83	27.48	5.17	16-37
BFI - Agreeableness	-.31*	-.34**	-.29*	.72	33.56	4.58	23-41
BFI - Conscientiousness	-.07	-.22	-.23	.86	31.55	6.82	16-45
BFI - Neuroticism	.02	-.17	.20	.83	23.92	5.14	14-34
BFI - Openness	.03	-.02	.15	.87	38.39	6.93	22-50
AQ - Physical Aggression	.49***	.54***	.37**	.75	22.24	5.89	11-38
AQ - Verbal Aggression	.36**	.33**	.31*	.72	1.80	3.56	5-20
AQ - Anger	.42***	.37**	.45***	.72	16.68	4.02	9-26
AQ - Hostility	.54***	.33**	.57***	.77	16.06	4.77	8-27
Self-esteem	.08	.14	-.07	.87	3.39	5.27	13-40
SOI-R - Behavior	.18	.13	.23	.75	5.62	2.76	3-17
SOI-R - Attitude	.41**	.40**	.29*	.87	11.09	6.64	3-25
SOI-R - Desire	.37**	.34**	.43***	.88	9.41	6.05	3-25
SOI-R - Total score	.44***	.42**	.42***	.84	26.12	12.26	9-55
MPP - Social Desirability	-.34**	-.10	-.35**	.76	21.67	4.64	12-33
MPP - Impression Management	.47***	-.03	.58***	.71	27.52	3.50	20-38

Note: n = 66; Ma = Machiavellianism; P = Psychopathy; N = Narcissism; α = Cronbach's alpha; M = Mean; SD = Standard deviation; MPP = Multidimensional Personality Profile; MMPI-PD = Minnesota Multiphasic Personality Inventory – Psychopathic Deviate; NPI = Narcissistic Personality Inventory; BFI = Big Five Inventory; AQ = Aggression Questionnaire; SOI-R = Sociosexual Orientation Inventory – Revised; * = $p < .05$; ** = $p < .01$; *** = $p < .001$;

Table 10 Descriptive statistics for the total sample, women, and men in Study 4.

Sample	Ma	N	P
Total sample (n = 974)	8.59±5.23 (.83, 4-28)	9.36±5.34 (.84, 4-28)	12.78±6.03 (.81, 4-28)
Women (n = 554)	7.71±4.76 (.84, 4-28)	8.43±4.96 (.84, 4-28)	12.40±6.10 (.84, 4-28)
Men (n = 420)	9.77±5.58 (.78, 4-28)	1.59±5.59 (.76, 4-28)	13.28±5.92 (.79, 4-28)

Note: Ma = Machiavellianism; P = Psychopathy; N = Narcissism; Values are mean±standard deviation (Cronbach's alpha, range)

Table 11 Goodness-of-fit statistics of confirmatory factor analytic and measurement invariance models in Study 4

Model and description	χ^2	df	CFI	TLI	RMSEA	ΔCFI	ΔTLI	ΔRMSEA	Ma	P	N
Total group	196.303	51	.956	.943	.054				-	-	-
Women	133.011	51	.953	.939	.054				-	-	-
Men	121.359	51	.953	.939	.057				-	-	-
Invariance models											
M0 Configural	255.136	102	.953	.939	.056				-	-	-
M1 Weak	26.863	111	.954	.945	.053	.001	.006	-.003	-	-	-
M2 Strong	288.541	120	.949	.944	.054	-.005	-.001	.001	.400***	.445***	.154*
M3 Strict	33.714	132	.942	.942	.056	-.007	-.002	.002	.397***	.442***	.154*
M4 Factor variances invariant	347.337	135	.939	.940	.057	-.003	-.002	.001	.438***	.481***	.150*
M5 Factor covariances invariant	35.252	138	.939	.940	.056	.000	.000	-.001	.438***	.481***	.151*

Note. χ^2 = Chi-square; df = degrees of freedom; CFI = comparative fit index; TLI = Tucker–Lewis index; RMSEA = root mean square error of approximation; Ma = standardized factor mean difference for Machiavellianism; P = standardized factor mean difference for Psychopathy; N = standardized factor mean difference for Narcissism. Positive standardized factor mean differences indicate higher scores in men. *** $p < .001$. * $p < .05$

Appendix

Italian version of the Dirty Dozen - DD-I

Troverai qui di seguito alcune affermazioni che possono o meno descrivere il tuo modo di essere, di pensare e di comportarti. Indica per ogni affermazione il tuo grado di accordo, ossia quanto l'affermazione ti sembra appropriata a descrivere la tua personalità, ricordando che 1 = fortemente in disaccordo e 7 = fortemente d'accordo

Non ci sono risposte "giuste" o "sbagliate". la migliore risposta è sempre quella che per prima ti viene in mente, in quanto è quella che ha la maggiore probabilità di essere la più sincera e quella che più si avvicina alla tua esperienza.

1.	Tendo a manipolare gli altri per ottenere ciò che voglio (*I tend to manipulate others to get my way)*	1	2	3	4	5	6	7
2.	Ho ingannato o mentito per ottenere ciò che volevo (*I have used deceit or lied to get my way*)	1	2	3	4	5	6	7
3.	Ho fatto ricorso all'adulazione per ottenere ciò che volevo (*I have used flattery to get my way*)	1	2	3	4	5	6	7
4.	Tendo a sfruttare gli altri per raggiungere i miei scopi (*I tend to exploit others towards my own end*)	1	2	3	4	5	6	7
5.	Tendo a non provare rimorso (*I tend to lack remorse*)	1	2	3	4	5	6	7
6.	Tendo a non preoccuparmi della moralità delle mie azioni (*I tend to be unconcerned with the morality of my actions*)	1	2	3	4	5	6	7
7.	Tendo a essere duro o insensibile (*I tend to be callous or insensitive*)	1	2	3	4	5	6	7
8.	Tendo a essere cinico (*I tend to be cynical*)	1	2	3	4	5	6	7
9.	Tendo a volere l'ammirazione degli altri (*I tend to want others to admire me*)	1	2	3	4	5	6	7
10.	Tendo ad esigere che gli altri mi prestino attenzione (*I tend to want others to pay attention to me)*	1	2	3	4	5	6	7
11.	Tendo a ricercare il prestigio ed un elevato status sociale (*I tend to seek prestige or status*)	1	2	3	4	5	6	7
12.	Tendo ad aspettarmi un trattamento speciale da parte degli altri *(I tend to expect special favors from others*)	1	2	3	4	5	6	7

Electronic Supplementary Materials for the paper *Does the Dark Triad Manifest Similarly in Men and Women? Measurement Invariance of the Dirty Dozen across Sex*

1. Power Analysis Study 1

The power analysis for confirmatory factor analysis (CFA) models of Study 1 was carried out using the procedure described in Muthén and Muthén (2002). The method relies on Monte Carlo simulations in which data are generated from a population with hypothesized parameter values. Ten thousand samples are drawn, and a model is estimated for each sample. Parameter values and standard errors are averaged over the samples and the following criteria are examined: parameter estimate bias, standard error bias, and coverage. In this case we followed the guidelines provided by the *Mplus User's Guide* (Muthén & Muthén, 1998–2010), Example 12.12, with the following settings for starting values:

- .80 for target loadings
- .00 for cross-loadings
- 2.5 for intercepts
- 1.00 for factor variances
- .50 for factor correlations in one group
- .20 for uniquenesses (residual variances)
- .00 for factor means

Muthén and Muthén (2002) suggest considering, as a first criterion, that parameter and standard error biases do not exceed 10% for any parameter in the model. The second criterion is that the standard error bias for the parameter for which power is being assessed does not exceed 5%. The third criterion is that coverage (i.e., the proportion of the replications where a 95% confidence interval covers the true parameter value) remains between .91 and .98. Once these three conditions are satisfied, the sample size is considered to keep power close to 0.80, a commonly accepted value for sufficient power.

We tested the power achieved by 5 different sample sizes: 100, 150, 200, 350 and 300. Results are reported in Table 1 and suggested that only Sample 3 afforded a sufficient statistical power to test the expected 3-correlated-factor CFA model.

Table 1 Parameter and standard error highest absolute bias and coverage for five different sample sizes to test the factor structure of the Dirty Dozen

Criteria	n= 100	n =150	n = 200	n = 250	n = 300
Parameter bias	6.78%	4.68%	3.60%	2.88%	2.54%
Standard error bias	18.39%	14.90%	13.21%	11.19%	9.51%
Coverage range	90.90-97.60%	92.30-97.30%	93.00-97.60%	93.40-97.50%	93.50-97.60%

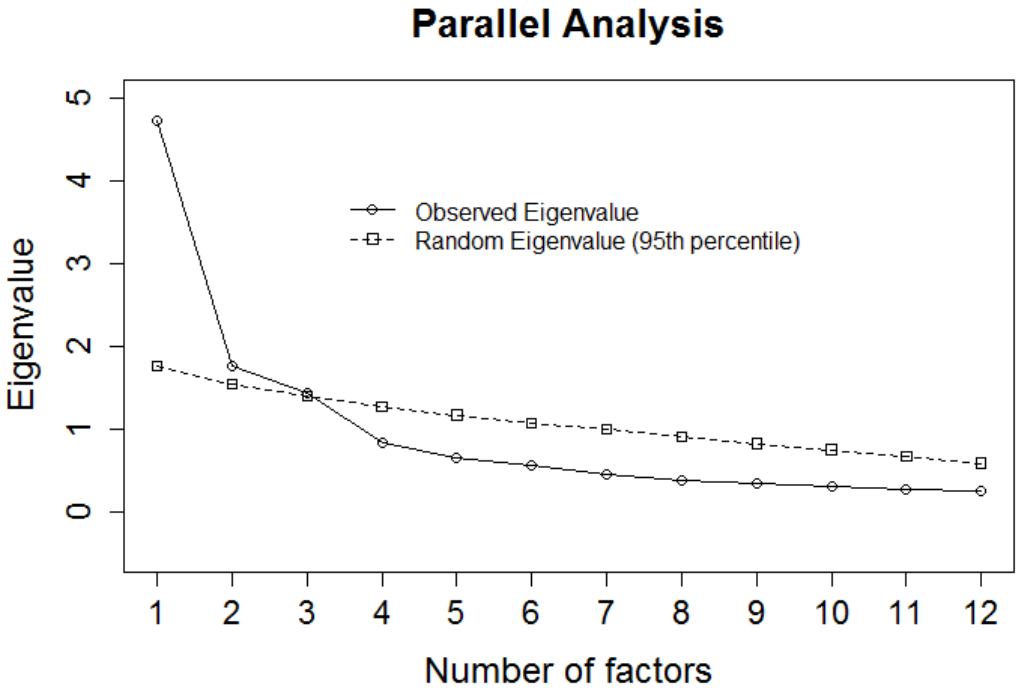
References

Muthén, B. & Muthén, L. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling, 4*, 599–620: doi: 10.1207/S15328007SEM0904_8

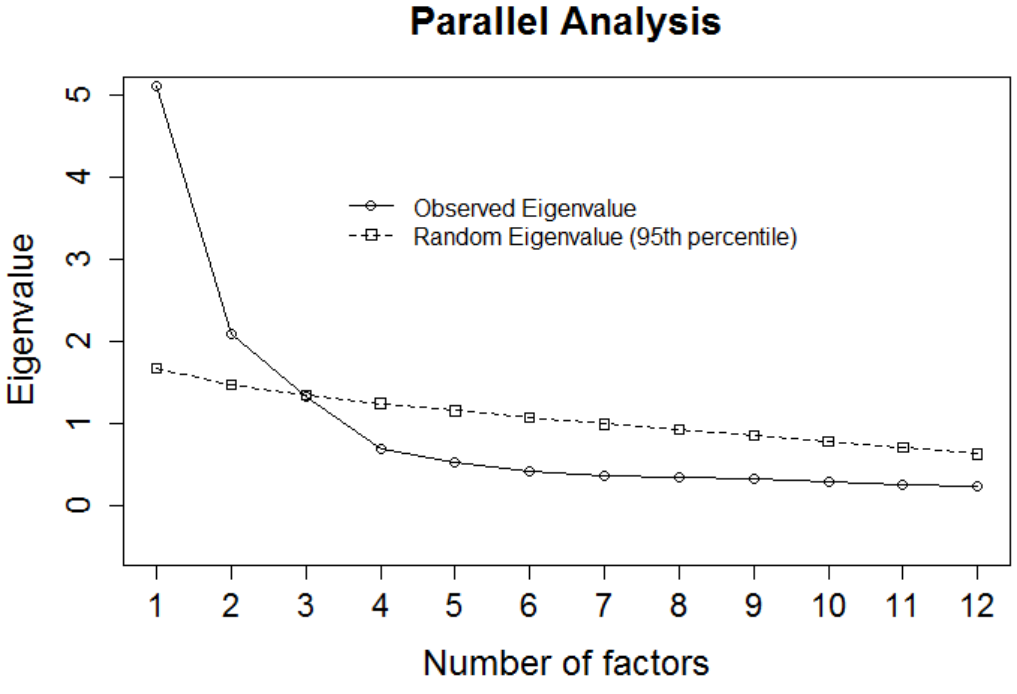
Muthén, L. K., & Muthén, B. (1998–2010). *Mplus user's guide*. Los Angeles, CA: Muthén & Muthén.

2. Analysis of the dimensionality of the item pool of the Italian Dirty Dozen in two independent samples of community participants (Study 1). PCA = Principal Component Analysis

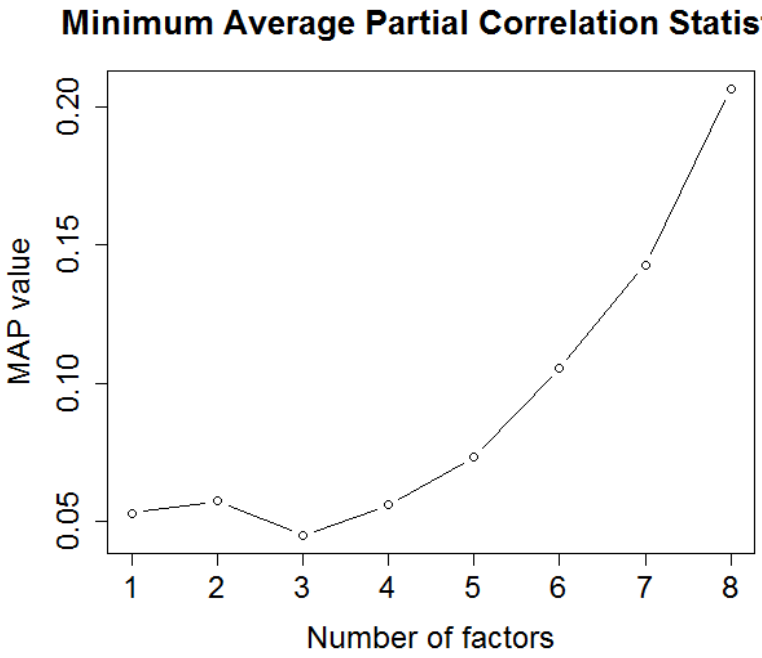
Sample 1 (*n*=102)



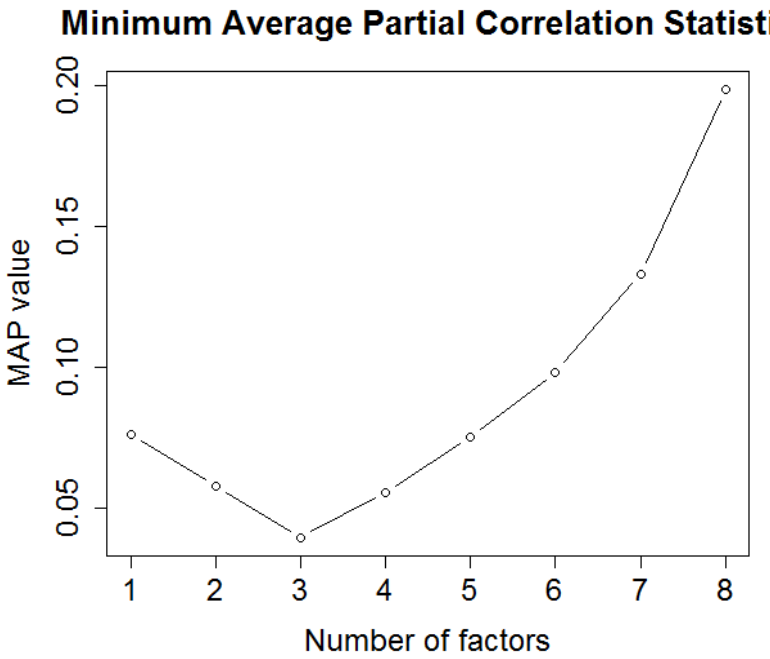
Sample 2 (*n*=128)



Sample 1 (*n*=102)

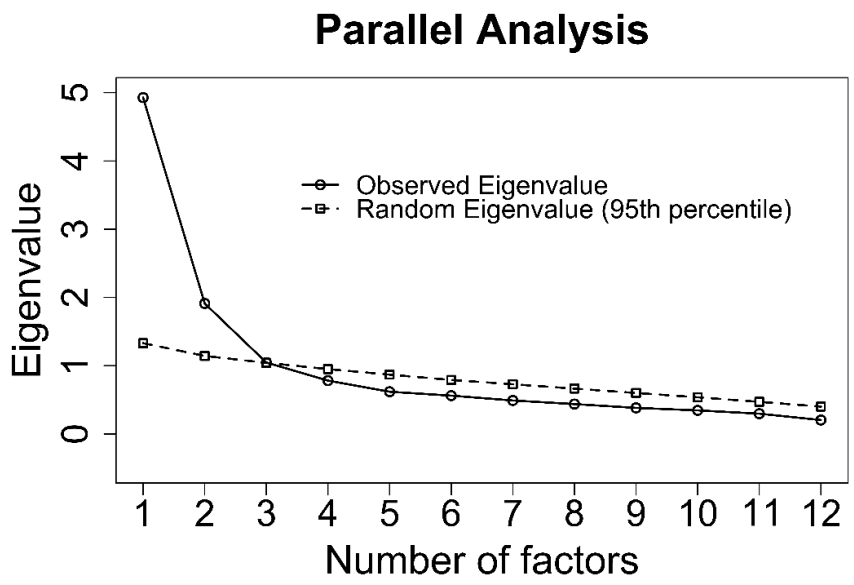


Sample 2 (*n*=128)

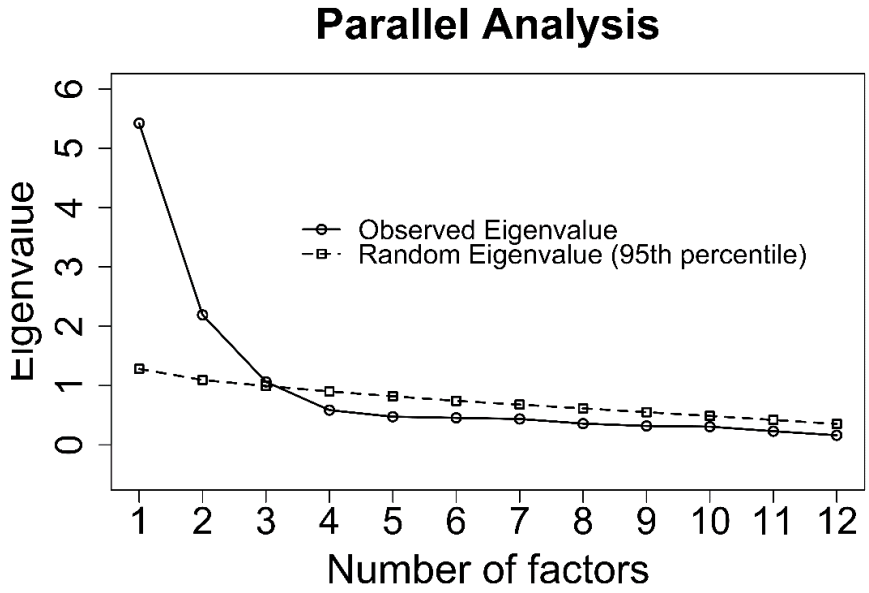


3. Dimensionality analyses and factor analyses for Italian Dirty Dozen in Study 2 (n = 164)

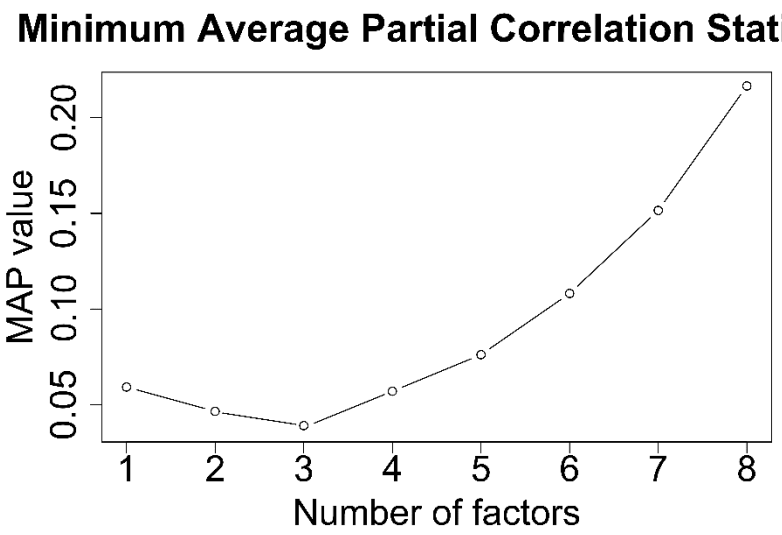
Time 1



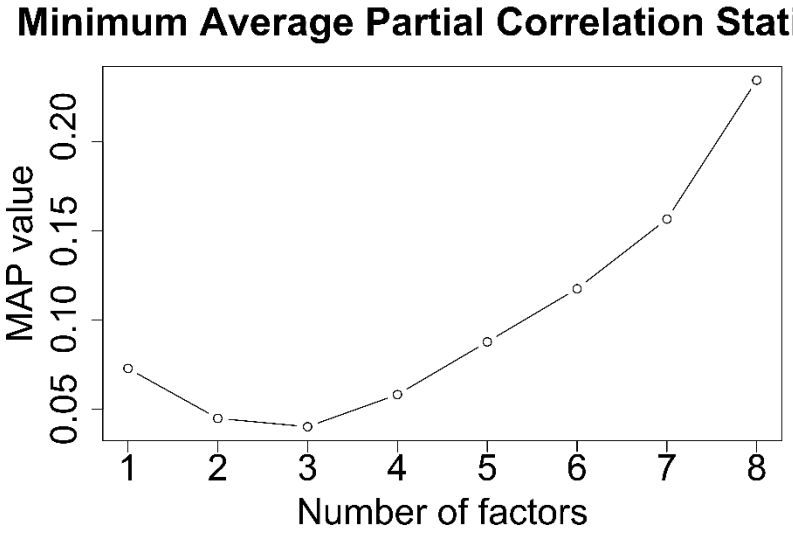
Time 2



Time 1


Minimum Average Partial Correlation Statistic
MAP value
0.05
0.10
0.15
0.20
1
2
3
4
5
6
7
8
Number of factors

Time 2


Minimum Average Partial Correlation Statistic
MAP value
0.05
0.10
0.15
0.20
1
2
3
4
5
6
7
8
Number of factors

4. Factor analysis and construct validity of the Italian version of the revised Sociosexual Orientation Inventory (SOI) in this study

As no formal Italian validation study of the SOI has been performed, we analyzed the factor structure and the construct validity of the inventory using data from Study 3. First, we carried out dimensionality analyses as those described in the manuscript for the DD. As shown in Figure 4.1, the scree-plot, the parallel analysis, and the MAP suggested that the optimal number of factors to extract was three.

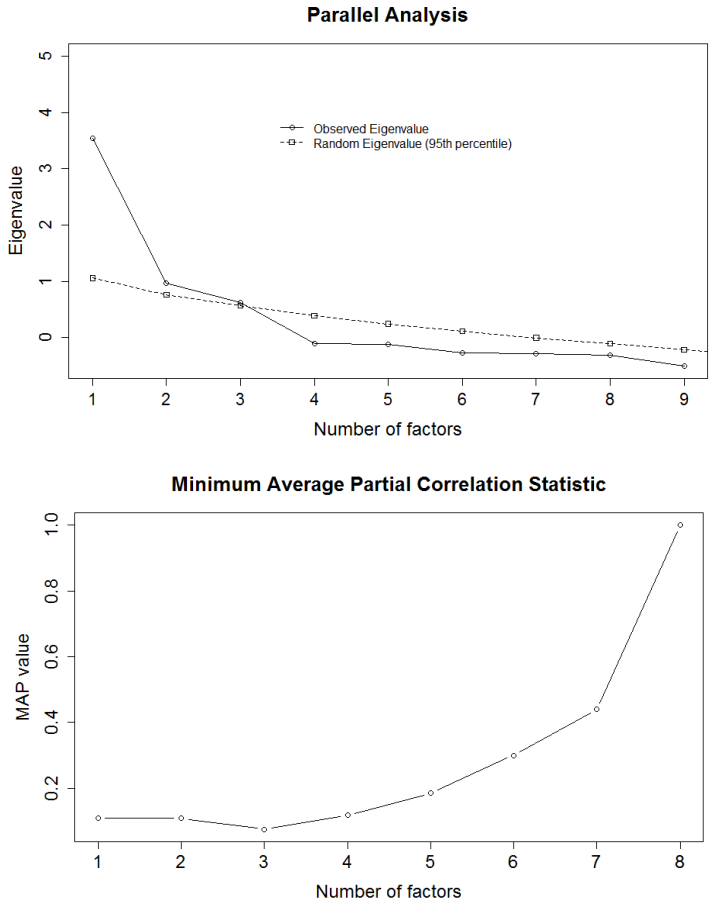


Figure 4.1 Results of the dimensionality analyses on the revised Sociosexual Orientation Inventory using data from Study 3

As the sample size (n = 66) did not afford sufficient statistical power to perform a confirmatory factor analysis, we carried out an exploratory factor analysis setting to the three the number of factors to be extracted- Promax rotation was used. This solution accounted for 67% of variance, and all item loaded on the expected factor, with minimal cross-loadings (Table 4.1).

Table 4.1 Results of the exploratory factor analysis on the Italian revised Sociosexual Orientation Inventory using data from Study 3

Item	F1	F2	F3
soi01	-.15	.19	.42
soi02	.16	.01	.73
soi03	-.03	.06	.91
soi04	.02	.73	.06
soi05	.02	.91	-.02
soi06	-.02	.80	.06
soi07	.84	-.02	.14
soi08	.86	-.06	.08
soi09	.81	.17	-.20
correlation with F2	.47		
correlation with F3	.28	.39	

Results of the correlation of the other measures employed in Study 3 supported the construct validity of the SOI total and subscale scores, as they were positively and significantly correlated with measures of the Dark Triad traits, of aggression, of extraversion, and of openness to experience, and negatively correlated with measures of agreeableness, conscientiousness, and social desirability (Table 4.2).

Table 4.2 Correlations of the Italian revised Sociosexual Orientation Inventory (SOI-R) scale scores with scores on measures of Machiavellianism, psychopathy, narcissism, personality, aggression, self-esteem, social desirability, and impression management (Study 3)

Scale	SOI-R - Total score	SOI-R - Behavior	SOI-R - Attitude	SOI-R - Desire
SOI-R - Behavior	.61**			
SOI-R - Attitude	.86**	.43**		
SOI-R - Desire	.80**	.30*	.45**	
DD - Machiavellism	.44**	.18	.41**	.37**
DD - Psychopathy	.41**	.13	.40**	.34**
DD - Narcissism	.42**	.23	.29*	.42**
AQ - Anger	.26*	-.03	.26*	.26*
AQ - Hostility	.37**	.19	.31*	.33**
AQ - Physical Aggression	.36**	.18	.35**	.26*
AQ - Verbal Aggression	.35**	.11	.24	.41**
BFI - Extraversion	.43**	.24	.28*	.46**
BFI - Agreeableness	-.25*	-.09	-.30*	-.15
BFI - Coscientiousness	-.43**	-.14	-.32**	-.45**
BFI - Neuroticism	-.09	.01	-.04	-.14
BFI - Openness	.21	.28*	.10	.20
MPP - Social Desirability	-.40**	-.25*	-.36**	-.29*
MPP - Impression Management	.22	.16	.21	.14
MPP - Machiavellism	.35**	-.08	.32**	.39**
NPI	.43**	.19	.36**	.39**
Self-esteem	-.06	.00	-.10	.00
MMPI-PD Total Score	.34**	.23	.24	.32**

Note: $n = 66$; DD = Dirty Dozen; AQ = Aggression Questionnaire; BFI = Big Five Inventory; MPP = Multidimensional Personality Profile; NPI = Narcissistic Personality Inventory; MMPI-PD = Minnesota Multiphasic Personality Inventory – Psychopathic Deviate; * = $p < .05$; ** = $p < .01$.

5. Socio-demographic differences between women and men in the sample of the paper

We tested whether the gender subgroups of the paper differed with respect to demographical variables. As shown in the table, some significant differences were found, but effect sizes were at best in the small range, suggesting that the comparisons of mean scores on the DD could have negligibly been biased (if ever) by the confounding effect of these variables.

Variable	Category	Women (n=554)	Men (n=420)	p	ES
Age (M±DS)		35.60±13.19	37.57±13.17	.021[a]	0.15[b]
Years of education (M±DS)		14.08±3.35	13.69±3.39	.065[a]	0.12[b]
Marital Status (proportion)	Single	.28	.23	.061[c]	.09[d]
	Married/Living together	.23	.18		
	Divorced/Separated	.04	.02		
	Widow/er	.01*	<.01*		
Occupation (proportion)	Unoccupied	.07*	.02*	<.001[c]	.19[d]
	Employed	.28*	.25*		
	Professional	.05	.06		
	Student	.15*	.08*		
	Retired	.01*	.02*		

Note: p = p-value for the statistical test; ES = effect size; M=mean; SD=standard deviation; [a]: independent sample t-test p-value; [b]: Cohen's *d*; [c]: chi-square test for the independence of categorical variables p-value; [d]:Cramer's *V*; *: column proportions statistically different ($p < .05$) after Bonferroni correction for multiple comparisons. Cohen's *d* is considered as negligible if $d <|.20|$, small if $|.20|<d<|.50|$, moderate if $|.50|<d<|.80|$, large if $d>|.80|$. Cramer's *V* is considered as negligible if $V <|.10|$, small if $|.10|<V<|.30|$, moderate if $|.30|<V<|.50|$, large if $V>|.50|$.

6. Measurement invariance models

Measurement invariance is usually tested with a sequence of models that impose equality constraints on the model parameters. Following Meredith (1964, 1993), the sequence of invariance testing begins with a model of *configural invariance* (M0), that is, with no invariance of any parameter estimates (i.e., all parameters are freely estimated), such that only similarity of the overall pattern of parameters is evaluated. This model tests whether the same factor structure is maintained across groups. Note that this model does not require any estimated parameters to be the same, hence it cannot be considered an actual invariance model. However, its fit has to be evaluated in order to provide both a test of the ability of the *a priori* model to fit the data in each group without invariance constraints and a baseline for comparing the other models that do impose equality constraints on the parameter estimates across groups.

The first step in invariance testing is to impose equality on factor loadings, i.e., specify a *weak* (or *scalar*) *invariance* model (M1). If identical items have statistically equivalent loadings, then the identical items show the same (if factor variances are fixed to 1 or constrained to be equal) or proportional (if variances are unequal) amount of increase between women and men for the same amount of increase on the latent factor (i.e., equality of scaling units; Millsap, 2011). This invariance is a prerequisite to comparisons of latent variances or relations among latent constructs. However, this model does not allow a test of differences in latent factor means, since mean differences based on latent constructs must be reflected in each of the individual items used to infer the latent constructs. It must then be shown that not only factor loadings, but also item intercepts (i.e., mean scores of individual items) are invariant over groups (*strong* or *scalar invariance* model, M2). If factor loadings and item intercepts are invariant over groups, then at all points along the factor continuum the same level of the latent factor results in statistically equivalent average scores on identical items between groups. This means that changes in the latent factor means can legitimately be interpreted as changes in the latent constructs. However, in models with freely estimated item intercepts and freely estimated latent means are not identified. Hence, the latent means are constrained to be zero in one group and freely estimated in the second group. This means that the freely estimated latent mean and its statistical significance reflect the differences between the two groups (Sörbom ,1974).

If one wants to compare (manifest) scale scores across groups, then an equality constraint must be posed also on item residual (or unique) variances (*strict measurement invariance* model, M3). This model assumes that same items have similar amounts of residual variance for both groups. We also tested models in which latent factor variances (M4) and covariances (M5) were constrained to be invariant. If latent factor variances are equal, in this case it would indicate that women and men used the same range on the factor continuum to report their levels of machiavellianism, psychopathy and narcissism and that the same items have equal levels of precision (reliability) across groups. If covariances are also invariant, the correlation between the same factor pairs for one group is statistically equivalent to the correlation between the same factors pairs for the other group.

References

Meredith, W. (1964). Notes on factorial invariance. *Psychometrika*, *29*(2), 177–185. doi:10.1007/BF02289699

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*(4), 525–543. doi:10.1007/BF02294825

Millsap, R. E. (2011). *Statistical approaches to measurement invariance*. New York, USA: Taylor & Francis Ltd.

Sörbom, D. (1974). A general method for studying differences in factor means and factor structure between groups. *British Journal of Mathematical and Statistical Psychology*, *27*(2), 229–239. doi:10.1111/j.2044-8317.1974.tb00543.x

7. Standardized coefficients from the configural invariance model

	Women (n = 554)				Men (n = 420)			
Items	Ma	P	N	RV	Ma	P	N	RV
DD01	.75			.44	.74			.46
DD02	.76			.42	.76			.43
DD03	.73			.47	.73			.46
DD04	.78			.39	.79			.38
DD05		.48		.77		.54		.71
DD06		.57		.68		.65		.57
DD07		.84		.30		.81		.34
DD08		.77		.41		.78		.40
DD09			.77	.41			.77	.42
DD10			.80	.37			.77	.41
DD11			.73	.47			.62	.61
DD12			.75	.44			.70	.51
Correlation with P	.51				.57			
Correlation with N	.59	.41			.59	.37		
Factor score determinacy	.93	.91	.93		.93	.91	.91	
Reliability (McDonald's ω)°	.81	.82	.71		.74	.79	.67	

Note. Ma = Machiavellianism; P = Psychopathy; N = Narcissism; I = intercept; RV = residual variance. All parameters are significant at $p < .001$; °: see SM for details.

8. Additional tables for the interpretation of the results of Study 3

Table 8.1 Correlations among DD-I scale scores across all studies

Correlation	Study 1 (Samples 1, 2, and 3)	Study 2 (T1 and T2)	Study 3	Study 4 (Total sample, Men, Women)
Ma with P	.46 , .45 , .46	.55 , .61	.50	.47 , .48 , .42
Ma with N	.42 , .50 , .50	.52 , .44	.66	.50 , .50 , .49
N with P	.31 , .26 , .33	.27 , .24	.24	.34 ,.31 , .35

Note. Ma = Machiavellianism; P = Psychopathy; N = Narcissism

Table 8.2 Correlations among DD scale scores in previously published studies.

Study	Ma-P	Ma-N	N-P
Aghababaei et al. / Personality and Individual Differences 67 (2014) 6–10	.40	.27	.29
Crysel et al. / Personality and Individual Differences 54 (2013) 35–40	.60	.51	.35
Czarna et al. Frontiers in Psychology, doi: 10.3389/fpsyg.2016.00445	.49	.39	.11
Fox, M.C. Rooney / Personality and Individual Differences 76 (2015) 161–165	.68	.50	.43
Furnham et al. / Personality and Individual Differences 67 (2014)	.56	.57	.32
114–121 (reports studies by others)	.52	.54	.31
	.51	.60	.34
	.60	.51	.35
Goodboy, M.M. Martin / Computers in Human Behavior 49 (2015) 1–4	.57	.52	.38
Jonason, P. K., et al. /Personality and Individual Differences 52 (2012), 449–453	.56	.57	.32
Jonason et al. / Personality and Individual Differences 78 (2015) 43–47	.58	.40	.30
Jonason, L. Krause / Personality and Individual Differences 55 (2013) 532–537	.62	.62	.37
Jonason, Webster / Psychological Assessment 22 (2010) 420–432 (Table 3)	.51	.34	.23
Maples et al. / Psychological Assessment 26 (2014) 326–331	.68	.56	.40
Lee et al. / European Journal of Personality 27 (2013) 169-184	.42	.32	.27
Spurk et al. / Social Psychological and Personality Science 7 (2016) 113-121	.61	.67	.55

Note. Ma = Machiavellianism; P = Psychopathy; N = Narcissism